Dear editor,

On behalf of the co-authors, I would like to thank you as well as the two anonymous reviewers for the constructive comments on our article. These comments and insights have been very helpful to improve the overall quality of the manuscript. They were all taken into account.

In particular, Referee #2 was concerned by the choice of a linear statistical model to relate pollutants and meteorology. We have addressed this concern by upgrading the statistical model to a Generalized Additive Model. The quality of the regression is slightly improved, but the general findings of the paper are unchanged, thereby consolidating their robustness. Referee #2 was also concerned by potential compensating effects on the various constituent contributing to the PM2.5 mix. A new section (3.2.2) where statistical models for each particulate matter constituent are explored was added to the manuscript. It demonstrates that there are no such compensating effects.

The main concern of Referee 1 regarded the choice of unique regional climate projections used to validate the methodology. We agree this is a valid point, however future simulations with the CHIMERE Chemistry-Transport Model but Regional Climate Model (RCM) other than the WRF IPSL-INERIS regional climate projections are not available. We are now working on the implementation of alternative climate forcing based on the selection of regional models identified in the present manuscript. In order to address the concern of Referee 1, we have better explained in the revised article why we have to rely on a calibration/testing ensemble based the present/future scenarios of the same RCM.

The detailed answers to the referee posted on the article discussion page address in more details these concerns and all the other points raised by both reviewers as well as all the changes made to the original submission. We hope you will consider that these changes demonstrate that our article deserve publication in ACP.

Thank you very much for your guidance in this reviewing process,

Best regards,

Vincent Lemaire.

Anonymous Referee #1

Received and published 25 November 2015

A methodology to "screen" regional climate projections in terms of their expected impact on future ozone and PM2.5 levels using a statistical model is presented. For PM2.5 the method only works for three out of eight regions in Europe which is a major concern, while for ozone it works better, six of eight regions.

Validation of the methodology was only done for one climate projection. Including one more of the regional climate projections in the validation of the methodology would make the paper more interesting and lend more confidence to the results presented.

It is our ambition to ultimately better understand the uncertainty that can be attributed to the climate forcing. To achieve this goal, we will force the CHIMERE CTM with alternative RCMs (in addition to the EuroCordex member of WRF-IPSL-INERIS that has been used here). However, because of the computational and storage demand required to retrieve such alternate forcing, we need to identify which RCM should be investigated in priority to offer an appropriate coverage of the range of uncertainty. It is the very purpose of the statistical ensemble exploration technique described in the present manuscript to define such priorities. However, at this stage alternative RCM forcings of the CHIMERE CTM are unfortunately not available to test the approach. Thus the only validation that we could include here was to test the statistical model on the basis of a future climate (2071-2100 in the RCP8.5), whereas it had been trained on a historical climate (1976-2005). The underlying hypothesis is that the historical range of meteorological parameters used to train the model will be exceeded in the future, therefore offering an appropriate testing dataset.

To better explain the difference between (i) ensemble of RCM forcing applied to CHIMERE and (ii) ensemble of RCM/CTM couples available in the literature, the following text has been added (page 3, line 15):

> "There are examples where more than two climate forcing are used, but then they are implemented with different CTMs, so that the uncertainties in the spread of RCM and CTMs is aggregated, thereby offering a poor understanding of the climate uncertainty. In addition, it should be noted that the choice of the climate driver is generally a matter of opportunity rather than an informed choice. These studies capture trends and variability but their coverage of uncertainty is not satisfactory in the climate change context."

To address the concern of the reviewer, the text has been modified page 28368 line 15 of the discussion paper (page 8, line 14 of the revised manuscript) in order to better explain why such a validation is not possible at present. The text now reads:

> "In order to evaluate the uncertainty related to climate change, the statistical models should be skillful for both pollutant concentrations over the historical period (training period) and in predictive mode.

Alternative RCM forcing of the CHIMERE CTM could be used to test the approach. Unfortunately, such alternatives are not available at this stage. The statistical ensemble exploration technique presented here will ultimately allow selecting the RCM that should be used in priority to cover the range of uncertainties in air quality and climate projections. When such simulations become available, we will be able to further test the skill of the statistical model. However, so far, the only validation that could be included here was to rely on a future time period as validation dataset. The underlying hypothesis is that the historical range of meteorological parameters used to train the model will be exceeded in the future, therefore offering an appropriate testing dataset. The results of this validation are presented in section 3.2."

The regions for which no robust relationships were found should be mentioned in the abstract.

Given that the statistical model only works for three regions for PM2.5 the phrase "The climate benefit for PM2.5 was confirmed" seems too strong. Also in view of what is presented in the introduction about the divergence of previous estimates of the climate benefit for PM2.5 in Europe.

The sign of the climate change impact on particulate matter is indeed still unclear in the literature. Therefore the term "confirmed" was removed from the abstract. We have made clearer the performances of the statistical model per region in the revised version, also including the regions where the skills of the statistical model were not satisfactory. The abstract now reads:

"In the three regions where the statistical model of the impact of climate change on PM2.5 offers satisfactory performances, we find a climate benefit (a decrease of PM2.5 concentrations under future climate) of -1.08 (± 0.21) $\mu g/m^3$, -1.03 (± 0.32) $\mu g/m^3$, -0.83 (± 0.14) $\mu g/m^3$, for respectively Eastern Europe, Mid Europe and Northern Italy. In the British Isles, Scandinavia, France, the Iberian Peninsula and the Mediterranean, the statistical model is not considered skillful enough to draw any conclusion for PM2.5."

"In Eastern Europe, France, the Iberian Peninsula, Mid Europe and Northern Italy, the statistical model of the impact of climate change on ozone was considered satisfactory and it confirms the climate penalty bearing upon ozone of 10.51 (± 3.06) $\mu g/m^3$, 11.70 (± 3.63) $\mu g/m^3$, 11.53 (± 1.55) $\mu g/m^3$, 9.86 (± 4.41) $\mu g/m^3$, 4.82 (± 1.79) $\mu g/m^3$, respectively. In the British Isles, Scandinavia and the Mediterranean, the skill of the statistical model was not considered robust enough to draw any conclusion for ozone pollution."

The two paragraphs above were added page 28362, line 14 (page 2, line 4 of the revised manuscript) to precise and rephrase the climate impact on particulate matter, according to the reviewer comment. The results section and the conclusion have also been rephrased.

This divergence in existing estimates of the climate impact on particulate matter is one of our main motivations to better document uncertainties in the climate impact on air quality. One of the important limitations is that almost all studies rely on a single source of climate projection. The climate uncertainty is therefore very poorly addressed. In this study, the use of a statistical model allows discussing the robustness of the climate effect on PM2.5 on the basis of the whole EuroCordex ensemble. Even if the statistical model only works for three regions out of eight for the PM2.5, the climate ensemble includes multiple global climate models driven by different regional climate models. Therefore the climate uncertainty is better covered by this study than by a projection based on a single climate source. This point has been made clearer in the introduction (page 28364, line 4 in the discussion paper and page 3, line 31 in the revised manuscript) that now reads:

> "This method allows selecting the members of the RCM ensemble that offer the widest range in terms of air quality response, somehow the "air quality sensitivity to climate change projections". These selected members should be used in priority in future air quality projections. A byproduct of our statistical air quality projections is that we explore an unprecedented range of climate uncertainty compared to the published literature that relies, at best, on two distinct climate forcings. The confidence we can have in these statistical projections is of course limited by the skill of the statistical model. Our approach of using a simplified air quality impact model but with a larger range of climate forcing can therefore be considered complementary with the more complex CTMs used with a limited number of climate forcings."

Although the focus of the paper is on the impact of climate change on air pollution it has been shown that projected European air pollution emission reductions has the potential to reduce both PM and ozone pollution in Europe to a large extent. This need to be mentioned in the introduction and the derived climate penalty/benefit should be contrasted to what could be achieved from emission reductions.

To address the reviewer comment, we have referred in more details to the literature in order to put in perspective the magnitude of climate and emission impacts. The following was added to the introduction (page 28363, line 6, of the discussion paper and page 2, line 27, of the revised manuscript):

> "There is therefore a concern that in the future, climate change could jeopardize the expected efficiency of pollution mitigation measures, even if the available studies indicate that if projected emission reductions are achieved they should exceed the magnitude of the climate penalty (Colette et al., 2013;Hedegaard et al., 2013)"

A quantitative comparison of our estimates of the impact of climate change with the impact of emission reduction strategies reported in the literature was also added in the conclusion (page 28379, line 13, of the discussion paper and page 20, line 18, of the revised manuscript):

> For PM2.5: "This impact of climate change on particulate pollution should be put in perspective with the magnitude of the change that is expected from the current air quality legislation. Such a comparison was performed by (Colette et al., 2013) who found (on average over Europe) a climate benefit by the middle of the century of the order of 0-1 µg/m3, therefore in line with our estimate but also much lower than the expected reduction of 7-8µg/m3 that they attributed to air quality policies."

> For Ozone: "It should be noted that when comparing the impact of climate change and emission reduction strategies, (Colette et al., 2013) found a climate penalty of the order of 2-3µg/m3 (which is broadly consistent with our results given that they focused on the middle of the century) that could be compensated with the expected magnitude of the reduction of 5-10µg/m3 brought about by air quality policies."

Section 2.2: the performance of the CTM for the historical period must be discussed. To what extent is the model capable of capturing observed variability in PM2.5 and ozone in a statistical sense? More information is needed here or reference to previous work documenting the performance.

The CTM used here has been extensively used and validated in the past, several references were added to the text in Section 2.2 (page 7 line 28), also focusing on the specific use of the model in the climate change context:

> "The Chemistry and Transport Model CHIMERE has been used in numerous studies: daily operational forecast (Rouïl et al., 2009), emission scenario evaluation (Cuvelier et al., 2007), evaluation in extreme events (Vautard et al., 2007), long term studies (Colette et al., 2011;Wilson et al., 2012;Colette et al., 2013) and inter-comparisons models and ensembles (Solazzo et al., 2012a;Solazzo et al., 2012b).

> The model performances depend on the setup but general features include a good representation of ozone daily maxima and an overestimation of night-time concentrations, leading to a small positive bias in average ozone (van Loon et al., 2007). Concerning particulate matter, similarly to most state-of-the-art CTMs, the CHIMERE model presents a systematic negative bias (Bessagnet et al., 2014). Regarding more specifically its implementation in the context of a future climate, evaluations of the CHIMERE model are presented in (Colette et al., 2013;Colette et al., 2015) and also (Watson et al., 2015) and (Lacressonniere et al., 2016)."

Europe should be included in the title. There is a need for language editing. Some suggestions are listed below but there is more to do to improve the readability.

We modified the title to include Europe and the text to improve the overall readability; we appreciate the reviewer specific recommendations. All the technical and language comments have been taken into account.

Technical/language comments

- p28362 l8: dataset from a deterministic
- p28363 l12: cost of such technique
- p28363 l16: amounts
- p28364 l5: dataset from a deterministic
- p28364 l19: such a
- p28364 l25: projections
- p28366 l3: decrease of concentrations
- p28366 l28: square based on using
- p28367 l3: have been
- p28369 l13: to obtain a strong climate signal and significant results.
- p28373 l20: can also give an
- p28374l10: latter
- p28374 l12: displayed in
- p28374 l21: in Fig. 3
- p28374 l25: in Fig. 3
- p28375 l1: compared
- p28375 l2: episodes
- p28375 l17: similar to the
- p28376 l1: a deeper
- p28376 l9: of the selected
- p28376 l20: axis
- p28377 l12: rise differs between
- p28377 l13: largest difference
- p28377 l18: which shows the lowest increase
- p28377l22: is associated
- p28377 l24: In Fig. S5
- p28378 l8: meteorological drivers
- p28378 l9: projections
- p28378 l10; climate change
- p28378 l15: meteorological drivers.

Anonymous Referee #2

Received and published: 17 November 2015

In this work we are introduced to a methodology, which is based on the use of a simple statistical model, which aims to confirm the impact of climate change impact on air quality and provide a range of uncertainty. The study is focusing on PM2.5 and ozone. While I understand the motivation for the development of such a handy tool, which will allow a fast and low-cost assessment of the impact of climate change on air quality, I am very skeptical on the methodology followed.

My major concern is the use of a very simplistic statistical model which tries to relate linearly PM2.5 with selected meteorological parameters. Considering that PM2.5 consist of a number of chemical species, which have various dependences on different meteorological variables (depending on their physical and chemical characteristics e.g. hygroscopicity, optical properties etc), I think it is not very responsible to assume a linear link between meteorology and PM2.5. My suggestion is that authors focus on ozone data only, and omit completely the analysis of PM2.5. Therefore, hereafter my comments will focus only on findings related to O3.

We understand the concern of the reviewer about the complexity of ozone chemistry and the fact that PM2.5 consists of numerous chemical species aggregated.

An important reason why the model we propose has skill is that we focus here on aggregated quantities (spatial averages, daily means), whereas non-linearities would have been much larger at high spatial resolution and temporal frequency. There are analogous findings in the study from (Thunis et al., 2015) who demonstrated that annual mean ozone and particulate matter responses to incremental emission changes were much more linear than previously thought. The following has been added in Section 2.1 to better explain this issue (page 6 line 21):

> "It should be noted that focusing on aggregated quantities greatly improves the skill of the statistical model that would struggle in capturing higher temporal frequency and spatial resolution. An analogy is presented in (Thunis et al., 2015) who demonstrated that annual mean ozone and particulate matter responses to incremental emission changes were much more linear than previously thought."

We also took the opportunity of this review to explore more sophisticated Generalized Additive Models as explained in more details when answering to the specific related comment below.

In order to address the Reviewer concern about possible compensating effects when using statistical regression for the PM2.5 mix of constituents, these more sophisticated statistical models were applied to each sub-constituent contributing to the PM2.5 mix and the corresponding results are presented in new Section 3.2.2 of the revised manuscript. That reads:

> “Because total PM2.5 is constituted by a mix of various aerosol species, there is a risk of compensation of opposite factors in the statistical model. In order to assess that risk, we developed such models for each individual PM constituent in the chemistry-transport model. The performances of these statistical models in terms of correlation for the historical (training) period or in predictive mode for the future period (testing) are presented in Figure 3.
>
> For all regions, the statistical models are not able to capture the variability of mineral dust. This is because the design of the statistical model is exclusively local (i.e. average concentrations over a given region are related to average meteorological variables over the same region), whereas most of the mineral dust over any European region is advected from the boundaries of the domain, in North Africa. It should be noted however, that except for the regions IP and MD, the dust represents only a small fraction of the PM concentrations (Figure 4). That could explain why the statistical model for PM2.5 performs poorly over IP and MD, but it will not undermine the confidence we can have in concluding about the robustness of the PM2.5 model for the region selected above: ME, EA and NI.
>
> All over Europe, primary particulate matter (PPM) is one of the smallest particulate matter fractions. Their variability is well captured by the statistical model for all the regions except SC. But because of their small abundance in that region, they should have a limited impact on the PM2.5 model performance.
>
> The sea salts are well reproduced by the statistical model for all the regions except NI and EA. These two regions have no maritime area, therefore sea-salt concentrations are lower and exclusively due to advection which, as a non-local factor, is not well captured by the statistical model.
>
> Ammonium (NH_4^+) aerosols are satisfactory captured by the statistical models for five regions out of eight including those selected for the overall PM2.5 model (ME, EA and NI).
>
> The organic aerosol fraction (ORG) is well reproduced over the historical period and the predictive skill is satisfactory (NRMSE around 0.7) for ME, EA and NI.
>
> The statistical models are efficient to reproduce the nitrate (NO_3^-) concentrations over the historical period for ME, EA, AL, MD, FR & BI regions but the predictive skills are only considered satisfactory for ME, EA, FR and NI, where nitrate constitutes a large fraction of PM2.5.
>
> Sulphate aerosols (SO_4^{2-}) are well represented by the statistical models for BI, EA and ME. The performances are low in the NI region, but sulphates constitute one of the smallest particulate matter fractions for that region.

This analysis of the skill of statistical models for each compound of the particulate matter mix confirms that there is no compensation of opposite factors in the selection of skillful models for total PM2.5 proposed in Section 3.2.1. The only cases were one of the particulate matter compound was not well captured by a statistical model, could be attributed to a low, and often non-local contribution of the relevant particulate matter constituent for the considered regions. We conclude that the selection of ME, EA and NI as regions where it is possible to build a statistical model of PM2.5 variability using Generalized Additive Models based on meteorological predictants would hold if the model had been built for each constituent of the particulate matter mix."

Because of the reasons explained above, we chose not to exclude PM2.5 from the analysis.

Throughout the manuscript, authors claim that with the use of their simple statistical model they can "conclude on the robustness of the climate impact on air quality". I think this statement exaggerates on what we can expect from a simple statistical model, the use of which, in my opinion, is to provide quick-looks or quick estimates, in a fast and low-cost manner. I would rather tend to consider as "conclusive", a result from a sophisticated numerical model, or even better, as authors also mention, results of several ensemble members.

Therefore, I would suggest the rephrasing of relevant sentences in the revised manuscript.

We rephrased the relevant sections, pointing out that such conclusions on the robustness (i) are limited by the skill of the statistical model, therefore only valid for selected regions and species, (ii) should be further tested by implementing full CTMs with the subset of RCMs identified as priorities.

We would like to emphasize that the supposed better quality of more sophisticated numerical models can at present only be reached at the cost of using a very limited number of regional climate projections (one in most studies, at best two in a couple of studies), therefore leaving a high risk in not offering an appropriate coverage of climate uncertainties. Our approach therefore offers a complementary assessment of such uncertainties, as now better explained in the introduction (page 3, line 31):

"This method allows selecting the members of the RCM ensemble that offer the widest range in terms of air quality response, which could be considered as air quality sensitivity to climate change projections. These selected members should be used in priority in future air quality projections. A byproduct of our statistical air quality projections is that we explore an unprecedented range of climate uncertainty compared to the published literature that relies, at best, on two distinct climate forcings. The confidence we can have in these statistical projections is of course limited by the skill of the statistical model. Our approach of using a simplified air quality impact model but with a larger range of climate forcing can therefore be considered complementary with the more complex CTMs used with a limited number of climate forcings."

I would strongly suggest the use of a more sophisticated statistical model, instead of the proposed linear model, to assess the impact of climate on O3. Especially if results prove to be different than those presented in the current analysis (i.e. if current findings are not reproducible), we can be sure that the current methodology suffers from several caveats –which we already known and authors already addressed - which cannot be ignored.

Building upon the advice of the referee, we have investigated the substitution of Linear Model by Generalized Additive Models (GAM) and also included a focus on each component of PM in Section 3.2.2.

We should however point out that none of the general conclusion of the original manuscript in terms of the climate penalty/benefit on ozone/PM is changed, and the RCM selected as a priority for further analysis are the same, thereby demonstrating the robustness of the approach itself.

"4.1 Fine particulate matter

In order to assess qualitatively the robustness of the evolution of regional climate variables having an impact on air quality, we first design a 2-D parameter space where the isopleths of statistically predicted pollutant concentrations are displayed (background of Figure 5). Then the distributions of historical and future meteorological variables as extracted from the regional climate projections are added to this parameter space. For each Regional Climate Projection, we show the average of the two driving meteorological variables as well as the 70th percentile of their 2D-density plot, i.e. the truncation at the 70^{th} quantile of their bi-histogram which means that 70% of the simulated days lie within the contour. Both historical and future climate projections (here for the RCP8.5 scenario and the 2071-2100 period) are displayed on the parameter space. The climate projections are all centered on the IPSL-CM5A-MR/WRF member so that only the distribution of the latter is shown for the historical period.

As pointed out in Table 1, the main meteorological drivers are the depth of the PBL and near surface temperature for the example of PM2.5 over Eastern Europe region displayed in Figure 5. The statistically modeled isopleths in the background of the figure show that PM2.5 concentration decrease when the depth of the PBL increases (x-axis), or when temperatures increase (y-axis). The interactions captured by the GAM exhibit the strong influence of high vertical stability events (with low surface temperature and PBL depth) in increasing PM2.5 concentrations. On the contrary, for high temperature ranges, the depth of the PBL becomes a less discriminating factor. The comparison of historical and future distributions shows that both meteorological drivers evolve significantly in statistical terms (Student t-test with Welch variant at the 95% confidence level based on annual mean). However, even though the PBL depth constitutes the most important meteorological driver for PM2.5, it does not evolve notably compared to the surface temperature in the future (Figure 5). Thus the largest increase of the secondary driver (surface temperature) leads to a decrease of PM2.5 concentrations. The largest and the smallest PM2.5 concentrations decrease are found for CSIRO-Mk3-6-0/RCA4 and MPI-ESM-LR/CCLM, respectively. But the overall spread of RCMs in terms of both the evolution of PBL depth and temperature is limited, suggesting that this climate benefit on particulate pollution is a robust feature. Those isopleths present the same characteristics for ME and NI regions (Supplementary information Figures S1, S4). The qualitative evolution represented in Figure 5 is further quantified by applying the GAM to the future meteorological variables in the regional climate projections. These results are represented by the probability density functions of the predicted concentrations of each GCM/RCM couple minus the estimated values for the historical simulation (e.g. 2071-2100 vs. 1976-2005, Figure 6). For EA and ME, the longer tail of the probability density function of MPI-ESM-LR/CCLM compared to the average of the models reflects that stronger pollution episodes will occur in the future even if the mean of the concentrations is lower than the average of the ensemble (Figure 6 for EA and Figure S2 for ME).

Besides the distribution, the ensemble mean and standard deviation of the estimated projected change in PM2.5 concentrations has been quantified (Table 3). All the selected regions depict a significant decrease of the PM2.5 concentrations across the multi-model proxy ensemble indicating that according to the GAM model, the climate benefit on particulate matter is a robust feature in these regions. The magnitude of the decrease depends on the region, its ensemble mean (± standard deviation) is -1.08 (± 0.21) $\mu g/m^3$, -1.03 (± 0.32) $\mu g/m^3$, -0.83 ± (0.14) $\mu g/m^3$, for respectively EA, ME and NI (Table 3).

In order to explain the differences in the response of individual RCM in the ensemble, we need to explore the historical meteorological variables probability density functions (PDF, Figure 7) and to compare them with the evolution of IPSL-CM5A-MR/WRF (Figure 5). The comparison of the historical distribution for the temperature reflects the stronger extremes of IPSL-CM5A-MR/WRF (e.g. colder than the others when it is cold). It is only for the NI region that IPSL-CM5A-MR/WRF lies in the mean of the ensemble. Concerning the PBL depth, the values are similar to the average of the ensemble for ME even if MPI-ESM-LR/RCA4 and EC-EARTH/RACMO2 present the largest values. IPSL-CM5A-MR/WRF has a thinner boundary layer for NI and a deeper than the average for EA but the differences are limited Figure 7).

It is for CSIRO-Mk3-6-0/RCA4 that we find the most important decrease of PM2.5 for the selected regions (Table 3). This is related to a larger temperature rise compared to the other models and a larger boundary layer height increase compared to the other member of the ensemble for these regions Figure 5). CanESM2/RCA4 and CSIRO-Mk3-6-0/RCA4 exhibit the same features for the ME region.

MPI-ESM-LR/CCLM presents the smallest decrease of PM2.5 for each of the selected regions (e.g. over ME is almost 3 times smaller than the largest decrease) except EA where CNRM-CM5-LR/RCA4 presents a smaller decrease (-0.77 $\mu g/m^3$ vs. -0.81 $\mu g/m^3$). As already mentioned above, the particular tails of the statistically modelled PM2.5 distributions for EA and ME indicate a larger contribution of large pollution episodes in the future for that RCM. But the historical distributions exhibit a larger boundary layer than the average models of the ensemble and a similar temperature. Thus, the low

PM2.5 concentration decrease is explained by the limited average evolution of the meteorological drivers as shown in Figure 5.

Overall we conclude that a climate benefit is identified for the PM2.5 for each of the selected regions. To the extent that the statistical model is skillful, as demonstrated in Section 3.2.1, this result is robust across the range of available climate forcings since the whole ensemble of regional climate projection present consistent features. The regional climate models that exhibit the largest and smallest responses are CanESM2/RCA4; CSIRO-Mk3-6-0/RCA4 and MPI-ESM-LR/CCLM, which should therefore be considered in priority for further evaluation using explicit deterministic projections involving full-frame regional climate and chemistry models.

4.2 Ozone peaks

For most of the selected regions (FR, IP, ME and NI,) the main drivers are the same (i.e. near surface temperature and short wave radiation). The isopleth in the background of Figure 5 show that temperature and short wave radiation have a similar impact on ozone peaks, except in the larger range of short wave radiation anomalies, where temperature becomes less discriminating. All the isopleths (Figure 5 for EA and Figures S1, S4 and S7 for ME, NI, FR and IP) exhibit an increase in the distribution of temperatures because the projected future is warmer than the historical period. According to the ozone peak concentrations predicted by the GAM (displayed in the background of Figure 5) these increases will lead to more ozone episodes. This trend appears for the entire models ensemble so that we can conclude with confidence that the climate penalty bearing upon ozone is a robust feature even if the specific distribution of some of the models stand out (CanESM2/RCA4; CNRM-CM5-LR/RCA4; CSIRO-Mk3-6-0/RCA4; IPSL-CM5A-MR/WRF).

The ozone increase of the ensemble reaches +10.51 (± 3.06) µg/m^3, +11.70 (± 3.63) µg/m^3, +11.53 (± 1.55) µg/m^3, +9.86 (± 4.41) µg/m^3, +4.82 (± 1.79) µg/m^3 for EA, FR, IP, ME and NI Table 3). These values confirm the statistically significant climate penalty (the mean is at least two times larger than the standard deviation). However, as already mentioned for Figure 5, we find minor differences among the models. The meteorological distributions are marginally different between the models of the ensemble: the summertime temperature predicted by IPSL-CM5A-MR/WRF has stronger extremes than the other models. Moreover, it is warmer than the ensemble in EA. Concerning incoming short wave radiation, IPSL-CM5A-MR/WRF lies in the average (Figure S3, S6, S9) except for the region EA where the amount of incoming radiation is the highest among the ensemble (Figure 7). Note that, only EC-EARTH/RACMO2 and MPI-ESM-LR/RCA4 exhibits lower values (around half of the average for MPI-ESM-LR/CCLM). The lower amount of summertime incoming short wave radiation for the couple MPI-ESM-LR/CCLM is relevant for all the selected regions.

The magnitude of the ozone rise differs between the models and the regions. Note that CanESM2/RCA4 exhibits the largest difference (i.e. around 1.5 times the ensemble mean) followed by CSIRO-Mk3-6-0/RCA4 for each selected regions. This is explained by the larger temperature increase during summertime which is the major driver, as identified by the statistical models, of ozone concentration. Note that the value is skyrocketing for the region ME, 5 times the value of IPSL-CM5A-MR/WRF which shows one of the lowest increases. CNRM-CM5-LR/RCA4 presents the lowest increase.

On the contrary, the lower increase of the summer temperature and sometimes a decrease of the incoming short wave radiation amount (e.g. IPSL-CM5A-MR/WRF in NI) lead to lower ozone concentration changes for IPSL-CM5A-MR/WRF and CNRM-CM5-LR/RCA4 for FR, IP, ME and NI (Table 3). Note the specific evolution for the region NI, where the IPSL-CM5A-MR/WRF model yields almost no increase of the ozone concentration compared to the other models while on the map of the differences in the deterministic model (Figure 1.f), the evolution was statistically significant. This absence of evolution reflects the limitation of the statistical models.

In figure S5, we can point out an outstanding pattern of the MPI-ESM-LR/CCLM distribution for the NI region with particularly large tails. The ozone rise would be more pronounced for the upper quantile which depicts more extreme ozone pollution episode (note that this was also the case for that model in terms of PM2.5 pollution).

Overall the climate penalty is confirmed even if some regional climate models stand out of the distribution, such as CanESM2/RCA4; CNRM-CM5-LR/RCA4 and CSIRO-Mk3-6-0/RCA4 which should therefore be considered for further deterministic projections."

We are not aware of previous studies highlighting caveats in the approach we propose, in particular because we focus on long term changes for aggregated indicators (regional averages and daily means), while it would indeed be more challenging to capture high-frequency changes at a monitor location for instance.

Another major concern comes from the resolution of the regional models used. A 50-km resolution is not recommended for impact studies. One cannot really expect much from the impact-assessment point-of-view in such a coarse resolution. Especially when higher resolution simulations are available in the community, what justifies the selection of 50 Km resolution model results? I would strongly suggest the use of 12 Km resolution simulations.

The appropriate model resolution differs with the climate impact being considered. As far as climate impacts on air quality are concerned, a higher resolution in the CTM is probably desirable to better address population exposure, even if it has been demonstrated that this improvement lies more in the refinement of emission inventories rather than meteorological (climate in our context) forcing (Valari and Menut, 2008).

Using a higher resolution climate forcing has been shown to improve the representation of extremes such as daily precipitation intensity (Jacob et al., 2014). However we should point out that air quality is not sensitive to precipitation extremes (triggering of low precipitation events, or blocking situations would be much more sensitive). In addition, again we limit the analysis to aggregated European regions, and (Kotlarski et al., 2014) demonstrated that for aggregated quantities over such subdomain, no apparent benefits of a finer grid are identified.

For air quality projection in a future climate context, 50km is the state of the art, none of the numerous papers published, even very recently, is using a higher resolution (Meleux et al., 2007;Langner et al., 2012a;Langner et al., 2012b;Manders et al., 2012;Colette et al., 2013;Hedegaard et al., 2013;Watson et al., 2015).

According to the reviewer comment, we add the following paragraph (page 7, line 15) to discuss the choice of such a resolution:

> "Whereas higher spatial resolution simulations are available in the EuroCordex ensemble, the 0.44 resolution were considered appropriate for air quality projections in agreement with other publications (Meleux et al., 2007;Langner et al., 2012a;Langner et al., 2012b;Manders et al., 2012;Colette et al., 2013;Hedegaard et al., 2013;Watson et al., 2015), and also because higher RCM resolution are not specifically performed to improve the climate features that are most sensitive for air quality purposes (temperature, solar radiation, stagnation events, triggering of low-intensity precipitation events etc.)."

However, if this is not possible, authors should definitely go into a very detailed description of evaluation and uncertainty issues of the climate data they used. Kotlarski et al 2014, Katragkou et al 2015 in Geosc Mod Dev and Garcia Diez et al, 2015 in Clim Dyn, provide details on the uncertainty issues and biases that stem from different EURO-CORDEX ensemble members. Authors could inform the readers in a concise way about the uncertainty (spread) associated with each meteorological variable used in the statistical model (temperature, precipitation, radiation) and how this may impact the results presented.

The model used to train the statistical model is the CTM CHIMERE forced by the RCM WRF that also contributed to the EuroCordex papers cited by the reviewer under the label "WRF-IPSL-INERIS". We added the following paragraph p 28369, line 3 of the manuscript (page 9, line 7 of the revised manuscript).

> "In the general EuroCordex evaluation, (Kotlarski et al., 2014) finds a good reproduction of the spatial temperature variability even if the models exhibit an underestimation of temperature during the winter in the north Eastern Europe. In addition to this general feature, the specificity of the WRF-IPSL-INERIS member is an overestimation of winter temperatures in the southeast. In terms of precipitations, most of the models exhibit a pronounced wet bias over most subdomains.
>
> When focusing on WRF members of the EuroCordex ensemble, (Katragkou et al., 2015) points out that the IPSL-INERIS member offers one of the best balance between precipitation and temperature skills."

The manuscript needs language editing (grammar, syntax, expression, punctuation). The title could be changed to become more informative, including more specific keywords such as “statistical model” instead of “proxies”, “ozone” instead of “air pollution”

We have modified the title to replace “proxies” by “statistical model”. All the technical comments have been taken into account.

Technical comments:

- Better to avoid using 3 dots for not complete lists, better use etc (e.g. page 28364, line 22)
- Abstract. Line 15. Replace “resp.” with respectively
- Introduction. Page 28363, line 8/9. “Hence the need to characterize…” the sentence is not complete.
- Page 28363, line 27. “The lack of multi-model approach in air quality and climate projections…” I think it should be in “air quality” or “climate/chemistry” projections and not “air quality and climate” projections, since there is no lack in climate projection ensembles.
- Page 28364, line 19. “have” replace with “has” Page 28364, line 1-4. Please rephrase.

Methodology:

- Page 28365, line 15. What do you mean by “phenomenological”?
  The word phenomenological was referring to chemical and physical processes. In order to avoid misunderstanding, we rephrased the sentence: “The choice of these meteorological variables is based on an analysis of the literature on the chemical and physical processes bearing upon air pollution and meteorology.”
- Page 28367, lines 21-23: “By using... over Europe”. Please rephrase.
  The rephrased sentence now reads: “By using deterministic climate and chemistry models from the global to the regional scale, they could produce long term air quality projections over Europe”.
- Page 28368, lines 24-26: This is a multi-model ensemble consisting of 12 members, 7 out of which are based on the same regional climate model. This implies that climatic information will be too much on the side of the RCA4 climate patterns. Authors could shortly discuss.
  There is indeed an imbalance in the matrix of GCM/RCM in the Eurocordex ensemble. This is an issue for the regional climate modelling community, and we have pointed that out in the revised manuscript page 9 line 17.
- Page 28369, lines 5-7. Could you please be more explicit?

  We rephrased to be more explicit (page 9, line 15): “Both studies are limited to the evaluation of RCM used with perfect boundary conditions (ERA-Interim forcing) and no published study has yet evaluated the various global/regional combinations.”
- Page 28370, line 16 (and Figure 1 caption). Authors should add in the text and in the figure caption the model suite out of which results are taken.
  A reference to the model suite has been added.
- Page 28370, line 13. You can skip “The” in “The Figure 1e corresponds to ...”.
  This change has been made.
- Page 28371, line 24. You don’t define SIA and SOA in text.
  According to the reviewer comment, we have defined Secondary Inorganic Aerosol and Secondary Organic Aerosol.
- Page 28372, lines 8- ..., please explain how you calculate R square/NRMSE for each subregion.

The following paragraph has been added to the text (page 12, line 7): "The NRMSE is defined as the root mean square error between statistically predicted and deterministically modelled concentrations changes aggregated by region and at daily temporal frequency, normalized by the standard deviation of the deterministic model. It allows describing the predictive power of a model, if the NRMSE is lower or equal to 1 then the model is a better predictor of the data than the mean (Thunis et al., 2012)."

- Page 28373, line 1-2. "This is because... long range transport of air pollution". Can you provide a reference for that? Why couldn't that be an impact of boundaries?

  Long range transport indeed includes pollution advected at hemispheric scale, therefore beyond the European boundaries covered by the model here. However, in the present context, boundary conditions are kept constant as explained section 2.2. Therefore, the long range transport we refer to here, is limited to interconnection within the European domain. It has been rephrased in the text (page 14, line 21).
- Page 28373, line 4-6: Where do we see the lower variability of temperature and incoming SW radiation? Do you suggest that Temperature and SW radiation are not relevant meteorological variables to explain O3 variability in south Europe? This does not make sense to me.

  We are not questioning the impact of temperature and incoming short wave radiation on ozone chemistry. However as a matter of fact they are not identified by the objective statistical analysis as the most discriminating variable because of their more limited range of variation in that region compared to other parts of Europe (standard deviation of 12.5 °C and 150 W/m² for MD; from 15 to 20°C and from 220 to 300 W/m² for the other regions). This has been added page 14, line 26.
- Section 4. This section could be more carefully written; with better structure and proper reference of uncertainty issues available in published literature (see general comments above).

  We have rewritten the section 4, following the general comments of the reviewer.
- Page 28374, line 17. How do you explain the fact that surface temperature increases significantly (did you check for significance? Better to avoid using if not properly tested), and PBL does not change notably? Any references that you could use?

  Both the increase of surface temperature and PBL change are statistically significant (Student t-test with Welch variant at the 95% confidence level based on the annual distribution aggregated per region) but the temperature increase is larger than the evolution of the PBL height. This is because surface temperature is not the only the driver of PBL depth (Menut et al., 2013). To address the reviewer comment the following text has been added page 16, line 1):

  > "The comparison of historical and future distributions shows that both meteorological drivers evolve significantly in statistical terms (Student t-test with Welch variant at the 95% confidence level based on annual mean). However, even though the PBL depth constitutes the most important meteorological driver for PM2.5, it does not evolve notably compared to the surface temperature in the future (Figure 5Figure)."

- Page 28377, line 5. I don't think you should use the word "slightly".

  The word "slightly" has been removed from the text.

Using ~~proxies~~statistical models to explore ensemble uncertainty in climate impact studies: the example of air pollution in Europe

V. E. P. Lemaire[1], A. Colette[1], and L. Menut[2]

[1]{Institut National de l'Environnement Industriel et des Risques (INERIS), Verneuil en Halatte, France}

[2]{Laboratoire de Météorologie Dynamique, UMR CNRS8539, Ecole Polytechnique, Ecole Normale Supérieure, Université P.M. Curie, Ecole Nationale des Ponts et Chaussées, Palaiseau, France}

Correspondence to: V. E. P. Lemaire (vincent.lemaire-etudiant@ineris.fr)

Abstract

Because of its sensitivity to unfavorable weather patterns, air pollution is sensitive to climate change so that, in the future, a climate penalty could jeopardize the expected efficiency of air pollution mitigation measures. A common method to assess the impact of climate on air quality consists in implementing chemistry-transport models forced by climate ~~projection~~projections. However, the computing cost of such ~~method~~methods requires optimizing ensemble exploration techniques.

By using a training dataset ~~of~~from a deterministic projection of climate and air quality over Europe, we identified the main meteorological drivers of air quality for 8 regions in Europe and developed ~~simple~~ statistical models that could be used to predict air pollutant concentrations. The evolution of the key climate variables driving either particulate or gaseous pollution allows ~~concluding on the robustness of the climate impact on air quality.~~ selecting the members of the EuroCordex ensemble of regional climate projections that should be used in priority for future air quality projections (CanESM2/RCA4; CNRM-CM5-LR/RCA4 and CSIRO-Mk3-6-0/RCA4 and MPI-ESM-LR/CCLM following the EuroCordex terminology).

~~The climate~~ After having tested the validity of the statistical model in predictive mode, we can provide ranges of uncertainty attributed to the spread of the regional climate projection ensemble by the end of the century (2071-2100) for the RCP8.5.

In the three regions where the statistical model of the impact of climate change on PM2.5 offers satisfactory performances, we find a climate benefit ~~for PM2.5 was confirmed -0.96 (± 0.18) µg/m^3,~~(a decrease of PM2.5 concentrations under future climate) of -1.~~00~~08 (± 0.~~37~~21) µg/m^3, -1.~~16 ± (~~03 (± 0.32) µg/m^3, -0.83 (± 0.~~23~~14) µg/m^3, for ~~resp.~~respectively Eastern Europe, Mid Europe and Northern Italy. In the British Isles, Scandinavia, France, the Iberian Peninsula and the Mediterranean, the statistical model is not considered skillful enough to draw any conclusion for ~~the~~PM2.5.

In Eastern Europe, France, the Iberian Peninsula, Mid Europe and Northern Italy~~ regions a~~, the statistical model of the impact of climate ~~penalty~~change on ozone was ~~identified~~considered satisfactory and it confirms the climate penalty bearing upon ozone of 10.~~11~~51 (± 3.~~22) µg/m^3, 8.23 (± 2.~~06) µg/m^3, ~~9.23~~11.70 (± 3.63) µg/m^3, 11.53 (± 1.~~13~~55) µg/m^3, ~~6~~9.86 (± 4.41~~ (± 2.14~~) µg/m^3, ~~7.43 (± 2.02~~4.82 (± 1.79) µg/m^3~~. This technique also allows selecting a subset of relevant regional climate model members that should be used in priority~~, respectively. In the British Isles, Scandinavia and the Mediterranean, the skill of the statistical model was not considered robust enough to draw any conclusion for ~~future deterministic projections~~ozone pollution.

1 Introduction

The main drivers of air pollution are (i) emission of primary pollutants and precursors of secondary pollutants, (ii) long-range transport, (iii) atmospheric chemistry and (iv) meteorology (Jacob and Winner, 2009). We can thus anticipate that air quality is sensitive to climate change taking as example the link between heat waves and large scale ozone episodes (Vautard et al., 2005)~~ as well as background changes.~~. But in addition to the direct impact of climate change on air pollution through the change in frequency and severity of synoptic conditions conducive to the accumulation of air pollutants we must also note that climate can have an impact on anthropogenic and biogenic _emission of pollutants and precursors (Langner et al., 2012b) as well as on changes in the global background of pollution, and therefore long range transport (Young et al., 2012). There is therefore a concern that in the future, climate change could jeopardize the expected efficiency of pollution mitigation measures~~ based on~~, even if the available studies indicate that if projected emission reductions~~. Hence the need to characterize and quantify uncertainties related to~~ are achieved they should exceed the ~~impact~~magnitude of the climate ~~change.~~ penalty (Colette et al., 2013;Hedegaard et al., 2013).

The most widespread technique used to assess the impact of climate change on air quality consists in implementing regional climate projections in Chemistry Transport Models (CTM) (Jacob and Winner, 2009). The computational cost of such ~~initiative~~technique is substantial given that it involves multi-annual global climate simulations, dynamical downscaling through regional climate simulations and ultimately CTM simulations. Besides the computational cost, it also raises technical difficulties in collecting, transferring and managing large ~~amount~~amounts of model data. Unlike many climate impact studies, CTM projections require Regional Climate Model fields in three dimensions and at high temporal frequency, whereas many regional climate modelling groups only store a few vertical levels in compliance with the CORDEX data archiving protocols. Altogether, these difficulties led to the use of a single source of climate projections in the majority of future air quality projections (Meleux et al., 2007;Katragkou et al., 2011;Jiménez-Guerrero et al., 2012;Langner et al., 2012b;Colette et al., 2013;Hedegaard et al., 2013;Varotsos et al., 2013;Colette et al., 2015) or two at most in published studies (Huszar et al., 2011;Juda-Rezler et al., 2012;Langner et al., 2012a;Manders et al., 2012~~;Colette et al., 2015~~)~~. And the choice of such a source was often~~. There are examples where more than two climate forcing are used, but then they are implemented with different CTMs, so that the uncertainties in the spread of RCM and CTMs is aggregated, thereby offering a poor understanding of the climate uncertainty. In addition, it should be noted that the choice of the climate driver is generally a matter of opportunity rather than an informed choice. These studies capture trends and variability but their ~~representation~~coverage of uncertainty is not satisfactory in the climate change context.~~ Moreover~~ This unsatisfactory handling of uncertainties is well illustrated by the divergence in ~~climate impact between two studies for~~ the ~~same pollutant supports again~~very sign of the ~~need of such ensemble approaches~~impact of climate change on particulate matter (e.g. (Lecœur et al., 2014) find a climate benefit for PM2.5 in Europe while

(Manders et al., 2012) suggest the opposite). Thus the lack of multi-model approach in air quality ~~and climate~~ projections is a serious caveat that needs to be tackled in order to comply with best practices in the field of climate impact research, where ensemble approaches is state of the art.

Hence, in order to assess the climate uncertainties on surface ozone and particulate matter over Europe in a changing climate, we developed ~~a new~~an alternative method which ~~avoids~~does not require forcing a CTM with an ensemble of climate models. It consists in ~~using~~developing a ~~simple~~ statistical model ~~applied~~fitted to ~~an~~a deterministic CTM simulation forced by a single RCM that can be subsequently applied to a larger ensemble of regional climate projections.

~~Using a training dataset~~ This method allows selecting the members of ~~deterministic projection~~the RCM ensemble that offer the widest range in terms of ~~climate and~~ air quality ~~over Europe, we identified the main meteorological drivers of~~ response, somehow the “air quality ~~for 8 regions in Europe and developed corresponding simple~~sensitivity to climate change projections”. These selected members should be used in priority in future air quality projections. A byproduct of our statistical ~~models that could be used to predict air pollutant concentrations. These~~ air quality projections is that we explore an unprecedented range of climate uncertainty compared to the published literature that relies, at best, on two distinct climate forcings. The confidence we can have in these statistical ~~models are subsequently applied to an ensemble of regional climate models (Jacob et al., 2014) to assess the robustness~~projections is of ~~the air quality projections. By discussing the evolution~~course limited by the skill of the ~~key climate variables of each member~~statistical model. Our approach of using a simplified air quality impact model but with a larger range of ~~the~~ climate ~~ensemble driving either particulate or gaseous pollution we can conclude on the robustness of~~ forcing can therefore be considered complementary with the more complex CTMs used with a limited number of climate ~~impact on air quality. Besides allowing a quantification of uncertainties, this technique also allows selecting a subset of relevant regional climate model members that should be used in priority for future ensemble deterministic projections.~~

forcings. The use of such a methodology is inspired from earlier work in the field of hydrology, where (Vano and Lettenmaier, 2014) have estimate future stream-flow by using a sensitivity-based approach which could be applied to generate ensemble simulations. Such a hybrid statistical and deterministic approach ~~have~~has also been used in the past in the field of air quality, but mostly for near-term and local forecasting, relying on statistical models of various complexity (i.e. Land Use Regression, Neural Network, Nonlinear regression, Generalized Additive Models~~…)~~ etc.) (Prybutok et al., 2000;Schlink et al., 2006;Slini et al., 2006). The most relevant example in the context of future air quality projection is that of (Lecœur et al., 2014), that use the technique of wind regime analogues, although they did not apply their approach to an ensemble of climate ~~projection~~projections.

This paper deals with all the steps needed to build the proxy of ensemble~~ and the results obtained.~~. First (Section 2) we present the methods and input data: the design of the statistical model of the air quality response to meteorological drivers is presented as well as the deterministic modelling framework used to create our training dataset. Section 3 focuses on results. The deterministic air quality projections are presented for ozone peaks and PM2.5 in Section 3.1. The selected statistical models for each region are evaluated in Section 3.2~~.~~ for ozone, PM2.5 and each sub-constituent of the particulate matter mix. The relevance of the statistical method to evaluate climate uncertainties and optimize ~~ensemble~~the exploration of the ensemble of climate projections is discussed in Section ~~4.~~4.

~~2 Development of statistical models of the air quality response to meteorological variability~~

~~2.1 Method~~

2 Methodology

2.1 Design

We consider ozone and PM2.5 as the main pollutants of interest for both purposes: public health (Dockery and Pope, 1994;Jerrett et al., 2009) and climate interactions (IPCC 2013). For both of them, ~~simple linear models have been developed using~~we investigated the best correlation that can be found for various European subregions using the following meteorological variables as predictants: near surface temperature (T2m), daily precipitation, incoming short wave radiation, planetary boundary layer (PBL) depth, surface wind (U10m) and specific humidity.

The choice of these meteorological variables is based on an analysis of the literature on ~~phenomenological links between~~the chemical and physical processes linking air pollution and meteorology. For PM2.5, turbulent mixing, often related to the depth of the planetary boundary layer, dominates (McGrath-Spangler et al., 2015). ~~Decrease~~A decrease of the PBL depth ~~is both related~~lead to either (i) an increase of the concentration of pollutants because the lower mixing volume (Jiménez-Guerrero et al., 2012) ~~and~~or (ii) a decrease of their concentrations because of their faster dry deposition ~~because of the vicinity of~~to surface receptors (Bessagnet et al., 2010). The wind plays also multiple roles for PM2.5. High wind speed favors the dilution of particulate matter (Jacob and Winner, 2009) but enhances sea-salt and dust mobilization (Lecœur and Seigneur, 2013). Precipitation is often reported as a major sink of PM2.5 through wet scavenging (Jacob and Winner, 2009). Water vapor~~, through specific humidity,~~ participates in aerosol formation during nucleation processes. Moreover, it can have an impact on the rates of certain ~~chemistry~~chemical reactions, ~~as well as~~similarly to temperature. The overall impact of temperature ~~impacts are~~on PM2.5 is difficult to isolate because of the ~~PM2.5~~ mix of components contributing to PM2.5 (organic, inorganic, dust, sea-salt…) and possible compensating effects. For instance, according to (Jacob and Winner, 2009), a temperature rise has opposite effects for ~~sulfate~~sulphate and nitrate (~~resp.~~respectively an increase and a decrease of concentrations). But for the overall PM2.5 mass, an increase in temperature will decrease the concentration as a result of higher volatility and subsequent higher aerosol to gas phase conversion (Megaritis et al., 2014). ~~In the case of the chemistry and transport model used in this study, CHIMERE (Menut et al., 2013), the volatile species in the gas and aerosol phases are assumed to be in chemical equilibrium. This thermodynamic equilibrium, computed by ISORROPIA (Fountoukis and Nenes, 2007), is driven by temperature and humidity and conditions the concentration of several aerosol species (ammonium, sodium, sulfate, nitrate and so on). Thus a major role of these variables is expected in this study.~~

~~The impact on~~As far as ozone ~~or its precursors are presented here. A~~is concerned, temperature ~~rise~~is expected to play a major role as it catalyzes atmospheric chemistry (Doherty et al., 2013). Moreover increasing temperature and solar radiation enhance isoprene emission which is a biogenic precursor of ozone (Langner et al., 2012b;Colette et al., 2013). Finally changing the amount of incoming short wave radiation will play a role on ~~the~~ozone photochemistry~~.~~

~~Indeed~~ , either by enhancing its photolysis by the hydroxyl radical in the presence of water vapor and short wave radiation ~~contributes both as a sink (water vapor & radiation leads to ozone photolysis) and a source (~~or by enhancing its production in the presence of photolysed nitrogen dioxide ~~photolysis produces ozone) of ozone~~ (Doherty et al., 2013). The impact of the PBL ~~effect~~depth on ozone varies with the meteorological conditions. Increasing the depth of the PBL dilutes ~~the~~ ozone concentrations, but it may also favors the ~~mixing of its precursors which leads~~dilution of nitrogen oxides close to the sources, therefore leading to an increase in ozone concentrations ~~increase~~in NOx saturated areas (Jacob and Winner, 2009). The amount of water vapor in the atmosphere mostly drives the abundance of the hydroxyl radical (OH). OH is ~~implied~~involved in ozone destruction through several processes (i.e. photolysis, HNO3 production) (Varotsos et al., 2013). It is also ~~implied~~involved in ozone production via the formation of NO2~~. Some VOC are oxidized by OH~~ and ~~form RO2 which reacts with NO to form NO2, a precursor of ozone~~radicals (Seinfeld and Pandis, 2008).

~~The design of~~ Starting from the above list of meteorological predictants, we aim to develop a statistical model ~~is deliberately limited to a simple bivariate linear least square based using two meteorological variables. In order to facilitate the geophysical interpretation, using meteorological variables instead~~ of ~~a linear combination of multiple variables (i.e. Prior Principal Component Analysis axes) is preferred. The main caveat is that it implies independence between meteorological variables.~~

~~While the skill of the statistical model could have improved by using a prior principal component analysis, a non linear model, or more than 2 predictors, we considered that remaining in a 2D physical parameter space was important for the purpose of the discussion as will be illustrated below. Hence the accuracy of the statistical proxy could be refined but we argue that our approach is satisfactory to assess uncertainty~~ozone and ~~optimize ensemble exploration.~~

~~Such a statistical model is built~~particulate matter for each of the eight European climatic regions ~~set~~defined in the PRUDENCE project (Christensen and Christensen, 2007). These regions are: British Isles (BI), Iberian Peninsula (IP), France (FR), Mid Europe (ME), Scandinavia (SC), Northen Italy (NI – referred to as the Alps in Climate studies but chiefly influenced by the polluted Po-Valley in the air quality context), Mediterranean (MD) and Eastern Europe (EA). For each of these regions, a spatial average of predictants (meteorological variables) and ~~predicted (~~pollutant concentrations~~)~~ values is taken. The statistical model is based on daily averages for all meteorological and air pollutant concentrations except ozone for which the daily maximum of 8-hr running means is used. The seasonality is removed by subtracting the average seasonal cycle over the historical period. It should be noted that focusing on aggregated quantities greatly improves the skill of the statistical model that would struggle in capture higher temporal frequency and spatial resolution. An analogy is presented in (Thunis et al., 2015) who demonstrated that annual mean ozone and particulate matter responses to incremental emission changes were much more linear than previously thought.

For each region and each pollutant, we first select the two most discriminating predictants by testing all the possible couple of meteorological variable and selecting those that reach the highest correlation. In a second stage we design the actual statistical model that consists of a Generalized Additive Model based on the two most discriminating perdictants (Wood, 2006).

It is to facilitate the geophysical interpretation that we use two meteorological variables instead of a linear combination of multiple variables (i.e. Prior Principal Component Analysis axes). Limiting their number to two also allows remaining in a 2D physical parameter space that supports the discussion as will be illustrated below.

2.2 Training and validation datasets

The datasets used to fit and test the statistical models are produced by the regional climate and air quality modelling framework presented in (Colette et al., 2013). By using ~~a full suite of models covering both~~deterministic climate and chemistry models from the global to the regional scale, they could produce long term air quality projections over Europe. The Earth System Model (ESM) which drives these simulations is the IPSL-CM5A-MR (Dufresne et al., 2013). The global data used in this study were produced for the Coupled Model Intercomparison Project Phase 5 initiative (CMIP5) (Taylor et al., 2012;Young et al., 2012). Then the climate data obtained by the ESM are dynamically downscaled ~~by~~with the regional climate model WRF (Skamarock et al., 2008). The spatial resolution is 0.44 degrees over Europe (Colette et al., 2013). These simulations were part of the low-resolution simulations performed within the framework of the European-Coordinated Regional Climate Downscaling Experiment program (EURO-CORDEX) (Jacob et al., 2014)~~. The~~

~~corresponding hindcasts were evaluated in~~. Whereas higher spatial resolution simulations are available in the EuroCordex ensemble, the 0.44 resolution were considered appropriate for air quality projections in agreement with other publications ~~(Kotlarski et al., 2014)~~(Meleux et al., 2007;Langner et al., 2012a;Langner et al., 2012b;Manders et al., 2012;Colette et al., 2013;Hedegaard et al., 2013;Watson et al., 2015)~~.~~, and also because higher RCM resolution are not specifically performed to improve the climate features that are most sensitive for air quality purposes (temperature, solar radiation, stagnation events, triggering of low-intensity precipitation events etc.). Finally the regional climate fields are used to drive the CTM CHIMERE (Menut et al., 2013), for the projection of air quality under changing climate. Since we are only interested in the effect of climate change, pollutant emissions remain constant at their level of 2010, as prescribed in the ECLIPSE-V4a dataset (Klimont et al., 2013). Similarly, chemical boundary conditions prescribed with the INCA model (Hauglustaine et al., 2014) as well as the land-use are also kept constant.

The Chemistry and Transport Model CHIMERE has been used in numerous studies: daily operational forecast (Rouïl et al., 2009), emission scenario evaluation (Cuvelier et al., 2007), evaluation in extreme events (Vautard et al., 2007), long term studies (Colette et al., 2011;Wilson et al., 2012;Colette et al., 2013) and inter-comparisons models and ensembles (Solazzo et al., 2012a;Solazzo et al., 2012b).

The model performances depend on the setup but general features include a good representation of ozone daily maxima and an overestimation of night-time concentrations, leading to a small positive bias in average ozone (van Loon et al., 2007). Concerning particulate matter, similarly to most state-of-the-art CTMs, the CHIMERE model presents a systematic negative bias (Bessagnet et al., 2014). Regarding more specifically its implementation in the context of a future climate, evaluations of the CHIMERE model are presented in (Colette et al., 2013;Colette et al., 2015) and also (Watson et al., 2015) and (Lacressonniere et al., 2016).

The training dataset used to build the statistical models ~~is~~consists of the historical air quality simulations (1976 to 2005), while ~~future~~ projections of air quality ~~and~~ under a future climate ~~projections~~(RCP8.5 2071-2100) will be used for testing purposes.

In order to evaluate the uncertainty related to climate change, the statistical models should be ~~efficient to reproduce the~~skillful for both pollutant concentrations over the historical period (training period) and ~~to predict them (testing period). We choose the~~ in predictive mode. Alternative RCM forcing of the CHIMERE CTM could be used to test the approach. Unfortunately, such alternatives are not available at this stage. The statistical ensemble exploration technique presented here will ultimately allow selecting the RCM that should be used in priority to cover the range of uncertainties in air quality and climate projections. When such simulations become available, we will be able to further test the skill of the statistical model. However, so far, the only validation that could be included here was to rely on a future time period as validation dataset~~ in order to challenge the statistical model trained over a given validity range that is expected to~~ . The underlying hypothesis is that the historical range of meteorological parameters used to train the model will be exceeded in the future~~. The different tests performed are explained~~, therefore offering an appropriate testing dataset. The results of this validation are presented in section 3.2.

2.3 ~~Regional climate projection ensemble~~Projection dataset

To evaluate the uncertainty related to the climate forcing, and identify the RCM that should be used in priority for future air quality projections, the statistical model of air quality is used in predictive mode using the regional climate projections performed in the framework of the EURO-CORDEX ~~program~~experiment (Jacob et al., 2014). The combinations of global/regional climate models used here are: CanESM2/RCA4; CSIRO-Mk3-6-0/RCA4; CNRM-CM5-LR/RCA4; EC-EARTH/RACMO2; EC-EARTH/RC4; GFDL-ESM2M/RCA4; IPSL-CM5A-MR/RCA4; IPSL-CM5A-MR/WRF; MIROC5/RCA4; MPI-ESM-LR/RCA4; MPI-ESM-LR/CCLM; NorESM1-M/RCA4 (see Jacob et al., 2014 for details on the model nomenclature).

The performances of the global models used to drive the regional projections have been evaluated in (Jury, 2012;Cattiaux et al., 2013)~~. The performances of the regional model driven by the ERA-Interim have been explored in~~ . In the general EuroCordex evaluation, ~~(Kotlarski et al., 2014)~~(Kotlarski et al., 2014)~~. No study has evaluated the bias of the global/regional combinations even if it could be relevant since the combination of a driving model and a regional model is not the simple addition of their bias. It is therefore safer to use such data to assess relative changes rather than absolute levels.~~ finds a good reproduction of the spatial temperature variability even if the models exhibit an underestimation of temperature during the winter in the north Eastern Europe. In addition to this general feature, the specificity of the WRF-IPSL-INERIS member is an overestimation of winter temperatures in the southeast. In terms of precipitations, most of the models exhibit a pronounced wet bias over most subdomains.

~~3 Results~~

~~In this part we studied the end (2071-2100) of the century, for one scenario (RCP8.5).~~When focusing on WRF members of the EuroCordex ensemble, (Katragkou et al., 2015) points out that the IPSL-INERIS member offers one of the best balance between precipitation and temperature skills. Both studies are limited to the evaluation of RCM used with perfect boundary conditions (ERA-Interim forcing) and no published study has yet evaluated the various global/regional combinations. It should also be noted that the ensemble is poorly balanced in terms of GCM/RCM combinations (see the larger weight of the RCA regional model which raise important question regarding the representativeness of the ensemble).

3 Development and validation of the statistical model

In this part we studied the end (2071-2100) of the century, for one scenario (RCP8.5) which is an energy-intensive scenario (van Vuuren et al., 2011). This 30 years period is chosen to be representative regardless of the inter-annual variability (Langner et al., 2012a).~~The RCP8.5 is a highly energy-intensive scenario (radiative forcing level equal to 8.5 W/m^2) (van Vuuren et al., 2011). We focus on the RCP8.5 and the end of the century to obtain the most significant results~~ We focus on the RCP8.5 and the end of the century on purpose to reach a strong climate signal.

3.1 ~~Climate and air~~Air quality projections

~~3.1.1 PM2.5~~

3.1.1 Fine particulate matter

Figure 1.a ~~represents~~shows the 30 years average PM2.5 concentrations over the historical period (1976 to 2005). Higher concentrations are modeled over European pollution hotspots: the Benelux, the Po Valley, Eastern Europe and large cities. A similar pattern is found in the future (RCP8.5 – average over the period 2071-2100) albeit with lower concentrations (Figure

1.b). The difference (future minus historical) is given in Figure 1.c where the statistical significance of the changes was represented by black points at each grid points and evaluated by a Student t-test with Welch variant at the 95% confidence level based on annual mean. The decrease is statistically significant over most of the domain.

Overall, we identify a climate benefit on particulate matter pollution similarly to (Colette et al., 2013;Lecœur et al., 2014)~~ (Lecœur et al., 2014)~~ but in opposition to (Manders et al., 2012). (Hedegaard et al., 2013) finds a decrease in high latitude and an increase in low latitude. The role of future precipitation projections and more efficient wet scavenging has often been pointed out to explain such a future evolution of particulate matter (Jacob and Winner, 2009). However, the lack of robustness in precipitation evolution over major European particulate pollution hotspots in regional climate models (Jacob et al., 2014) challenges the confidence we can have in single model air quality and climate projection, supporting again the need for ensemble approaches.

3.1.2 Ozone peaks

Figure 1.d represents the summer (JJA) average ozone daily maximum concentrations over the historical period (1976 to 2005). A North-South gradient appears with lower concentration in the North and higher concentration fields over the Mediterranean Sea. ~~The ~~Figure 1.e corresponds to the summer average ozone projection of the RCP8.5 at the end of the century (2071-2100~~).~~) predicted by the model suite presented in Section 2.2. A similar pattern is found, with higher concentrations in the southern part of the domain~~.~~ (Figure 1.e). The map of the difference~~, Figure 1.f,~~ (RCP8.5 - actual~~)~~), Figure 1.f, indicates an increase of ozone concentrations over Eastern Europe, Mediterranean land surfaces, and North Africa and a decrease over British Isles and Scandinavia. Most of the changes are statistically significant except over West Europe. This concentration rise is frequently associated to an increase of temperature in the literature (Meleux et al., 2007;Katragkou et al., 2011), see Section ~~2~~2.1 above for ~~an~~a review of physical and chemical processes underlying this association.

Following (Langner et al., 2012b;Manders et al., 2012;Colette et al., 2013;Colette et al., 2015) ~~we~~these result confirm~~ overall~~ the fact that climate change constitutes a penalty for surface ozone in Europe.

3.2 Statistical models

Here we introduce the statistical models trained over the historical period~~.~~ and their evaluation over the future testing period. First we discuss the~~ expected~~ impact of key meteorological processes on pollutants concentration on the basis of the model correlation and put our results in perspective with the key driving factors reported in the literature. Then we evaluate the performance of statistical models over the future period in order to discard regions and pollutants where the skill of the statistical model is too small to draw robust conclusions on the uncertainties of projections.

~~3.2.1 PM2.5~~

3.2.1 Fine particulate matter

The skill and predictors for ~~bivariate linear least square statistical~~generalized additive models fitted for each region are given in Table 1. The depth of the planetary boundary layer is identify as the major meteorological driver for PM2.5 which is a different finding compared to (Megaritis et al., 2014) who report a smaller impact for the PBL depth. Near surface temperature is often selected as second predictor. The wind is pointed out as a relevant predictor twice ~~and~~but only for coastal regions (~~resp.~~respectively BI and MD) where ~~the ~~sea-salt ~~dominates~~is important. Last, precipitation is selected only once and as 2^{nd} variable for the Iberian Peninsula (IP). It could be partly due to our choice of ~~a linear correlation~~statistical model whereas a logical regression would have been more efficient given that PM correlations are sensitive to the presence/absence of precipitation rather than their intensity. It is difficult to assess objectively whether the larger role of temperature than precipitation in our findings is an artifact related to the design of the statistical model. The importance of precipitation in the impact of climate change on particulate pollution is often speculated in the literature, with little quantitative evidence. The ~~bilinear~~statistical model used here~~ is simplistic, but it~~ offers an objective quantification of that role. It should be added that the importance of temperature is well supported by the volatilization process for Secondary Inorganic Aerosol and Secondary Organic Aerosol. Moreover in the CTM CHIMERE, the volatile species in the gas and aerosol phases are assumed to be in chemical equilibrium. This thermodynamic equilibrium, computed by ISORROPIA (Fountoukis and Nenes, 2007), is driven by temperature and humidity and conditions the concentration of several aerosol species (ammonium, sodium, ~~SIA and SOA.~~sulphate, nitrate and so on). This feature could explain the major role of temperature. It is also supported by the pattern of projected PM2.5 change, which is spatially correlated with present-day PM2.5 concentration. This spatial correlation suggests an impact of a uniform driver which points towards temperature rather than precipitation change that exhibits a strong north-south gradient in Europe.

Then the predictive skill of these models is tested over the period 2071-2100 by computing the Normalized Root Mean Squared Error (NRMSE) between the statistically predicted PM2.5 (concentrations estimated with the statistical models), and the results of the deterministic regional air quality and climate modelling suite presented in section 2.2~~.~~ for 2071-2100. The NRMSE is defined as the root mean square error between statistically predicted and deterministically modelled concentrations changes aggregated by region and at daily temporal frequency, normalized by the standard deviation of the deterministic model. It allows describing the predictive power of a model, if the NRMSE is lower or equal to 1 then the model is a better predictor of the data than the data mean (Thunis et al., 2012).

Figure 2 shows, for each region, the scatter between r-squared over the historical period and the NRMSE in predictive mode for the RCP8.5 at the end of the century.~~ The NRMSE is equal to the RMSE divided by the standard deviation of the reference: air quality projection (2071-2100). It allows describing the predictive power of a model, if the result is "less or equal to 1" then the model is a better predictor of the data than the data mean (Thunis et al., 2012).~~ We expect regions where the correlation over the historical period is low to be poorly captured by the statistical model in the future. The fact that the good correlation for EA and ME are associated with a NRMSE around 0.~~7~~6 in the future indicates either that the main meteorological drivers in the future will remain within their range of validity or that extrapolation is a viable approximation.

This feature gives confidence in using statistical models for these regions in predictive mode. For the NI region, the NRMSE is acceptable (below 0.~~85~~8) even if the r-squared is low.

Considering that the model skill was satisfactory for the EA, ME and NI regions, we decided to focus on these regions for the uncertainty assessment in the remainder of this paper. The fine particulate matter concentrations have been poorly captured for the region BI, SC, FR, IP and MD. The associated bad NRMSE are explained by the poor performances of model over the historical. They are thus excluded from the uncertainty assessment.

3.2.2 Particulate matter composition

Because total PM2.5 is constituted by a mix of various aerosol species, there is a risk of compensation of opposite factors in the statistical model. In order to assess that risk, we developed such models for each individual PM constituent in the chemistry-transport model. The performances of these statistical models in terms of correlation for the historical (training) period or in predictive mode for the future period (testing) are presented in Figure 3.

For all regions, the statistical models are not able to capture the variability of mineral dust. This is because the design of the statistical model is exclusively local (i.e. average concentrations over a given region are related to average meteorological variables over the same region), whereas most of the mineral dust over any European region is advected from the boundaries of the domain, in North Africa. It should be noted however, that except for the regions IP and MD, the dust represents only a small fraction of the PM concentrations (Figure 4). That could explain why the statistical model for PM2.5 performs poorly over IP and MD, but it will not undermine the confidence we can have in concluding about the robustness of the PM2.5 model for the region selected above: ME, EA and NI.

All over Europe, primary particulate matter (PPM) is one of the smallest particulate matter fractions. Their variability is well captured by the statistical model for all the regions except SC. But because of their small abundance in that region, they should have a limited impact on the PM2.5 model performance.

The sea salts are well reproduced by the statistical model for all the regions except NI and EA. These two regions have no maritime area, therefore sea-salt concentrations are lower and exclusively due to advection which, as a non-local factor, is not well captured by the statistical model.

Ammonium (NH_4^+) aerosols are satisfactory captured by the statistical models for five regions out of eight including those selected for the overall PM2.5 model (ME, EA and NI).

The organic aerosol fraction (ORG) is well reproduced over the historical period and the predictive skill is satisfactory (NRMSE around 0.7) for ME, EA and NI.

The statistical models are efficient to reproduce the nitrate (NO_3^-) concentrations over the historical period for ME, EA, AL, MD, FR & BI regions but the predictive skills are only considered satisfactory for ME, EA, FR and NI, where nitrate constitutes a large fraction of PM2.5.

Sulphate aerosols (SO_4^{2-}) are well represented by the statistical models for BI, EA and ME. The performances are low in the NI region, but sulphates constitute one of the smallest particulate matter fractions for that region.

This analysis of the skill of statistical models for each compound of the particulate matter mix confirms that there is no compensation of opposite factors in the selection of skillful models for total PM2.5 proposed in Section 3.2.1. The only cases were one of the particulate matter compound was not well captured by a statistical model, could be attributed to a low, and often non-local contribution of the relevant particulate matter constituent for the considered regions. We conclude that the selection of ME, EA and NI as regions where it is possible to build a statistical model of PM2.5 variability using Generalized Additive Models based on meteorological predictants would hold if the model had been built for each constituent of the particulate matter mix.

~~3.2.2~~3.2.3 Ozone peaks

For summertime ozone peaks, as expected, near surface temperature and incoming short wave radiation are identified as the two main meteorological drivers for most regions ~~(Cf.~~ (Table 2). Concerning the region EA, the drivers which give the best results are near surface temperature and specific humidity. Nevertheless, when using specific humidity as second predictor, the statistical model is overfitted and has a low predictive skill (NRMSE=0.9). Thus the use of short wave radiation as second predictor appears much more robust (NRMSE=0.6) even if the R^2 is lower. The skill of the statistical model is very low over the British Isles and Scandinavia. This is because ozone pollution in these regions is largely influenced by ~~the~~ non-local contributions (long range transport of air pollution~~. It is therefore poorly correlated with the local variability of meteorological variables.~~). The poor performances of the statistical model over the Mediterranean region are more surprising. The lower variability of temperature and incoming shortwave radiation in this region ~~makes them less relevant to explain~~compared to other parts of Europe (standard deviation of 12.5 °C and 150 W/m² for

MD; from 15 to 20°C and from 220 to 300 W/m² for the other regions) makes them less relevant as statistical predictants of ozone concentrations.

~~The linear~~We conclude that the generalized additive models that ~~are ultimately~~can be considered efficient enough in terms of correlation to capture the ozone ~~concentrations~~variability over the historical period are those of the following regions: EA, FR, IP, ME and NI.

This selection is further supported by ~~applying the same approach as for PM2.5 to evaluate~~investigating the predictive skill of the models~~.~~ assessed by computing their NRMSE against deterministic CTM simulations available for a future period. The regions mentioned above where the correlation of the statistical model is low (BI, SC and MD) ~~stand out on this graph (Cf.~~ also exhibit a large NRMSE (Figure 2). So that, only the regions~~:~~ EA, FR, IP, ME and NI are selected for the ~~uncertainty assessment in the~~ remainder of this paper.

~~4 Evaluation~~Exploring the ensemble of ~~the uncertainty~~

4 ~~To evaluate the robustness of the air quality and~~ climate ~~projection we use~~projections with the statistical model

The statistical models introduced in Section 2~~-~~, developed in Section 3 and tested in Section 3.2 are applied here to the ensemble of regional climate projections presented in Section 2.3 to develop a proxy of ensemble of air quality and climate projections for each selected region. This proxy of ensemble will be used to identify the subset of regional climate projections that should be used in priority in the deterministic modelling suite, but it can also ~~given~~give an indication on the robustness of the climate impact on air quality ~~by comparing~~where the ~~evolution~~skill of ~~key climate drivers~~the statistical model is considered satisfactory.

~~4.1 PM2.5~~

4.1 Fine particulate matter

In order to assess qualitatively the robustness of the evolution of regional climate variables having an impact on air quality, we first design a 2-D parameter space where the isopleths of statistically predicted pollutant concentrations are displayed (background of Figure ~~3~~5). Then the distributions of historical and future meteorological variables as extracted from the regional climate projections are added to this parameter space. For each Regional Climate Projection, we show the average of the two driving meteorological variables as well as the 70th percentile of their 2D-density plot, i.e. the truncation at the 70th quantile of their bi-histogram which means that 70% of the simulated days lie within the contour. Both historical and future climate projections (here for the RCP8.5 scenario and the 2071-2100 period) are displayed on the parameter space. The climate projections are all centered on the IPSL-CM5A-MR/WRF member so that only the distribution of the ~~later stands out~~latter is shown for the historical period.

As pointed out in Table 1, the main meteorological drivers are the depth of the PBL and near surface temperature for the example of PM2.5 over Eastern Europe region displayed ~~on~~in Figure ~~3~~5. The statistically modeled isopleths in the background of the figure show that PM2.5 concentration decrease when the depth of the PBL increases (x-axis), or when temperatures increase (y-axis). ~~The comparison of historical and future distributions shows that~~The interactions captured by the GAM exhibit the strong influence of high vertical stability events (with low surface temperature and PBL depth) in increasing PM2.5 concentrations. On the contrary, for high temperature ranges, the depth of the PBL becomes a less discriminating factor. The comparison of historical and future distributions shows that both meteorological drivers evolve significantly in statistical terms (Student t-test with Welch variant at the 95% confidence level based on annual mean). However, even though the PBL depth constitutes the most important meteorological driver for PM2.5, it does not evolve notably compared to the surface temperature in the future ~~(Cf.~~ (Figure ~~3~~5). ~~On~~Thus the ~~contrary,~~largest increase of the secondary driver (surface temperature) ~~increases significantly, leading~~leads to a decrease of PM2.5 concentrations. The ~~small~~largest and the smallest PM2.5 concentrations decrease are found for CSIRO-Mk3-6-0/RCA4 and MPI-ESM-LR/CCLM, respectively. But the overall spread of RCMs in terms of both the evolution of PBL depth and temperature ~~suggests~~is limited, suggesting that this climate benefit on particulate pollution is a robust feature. ~~However, on Figure 3, CSIRO-Mk3-6-0/RCA4 and MPI-ESM-LR/CCLM present respectively the largest and the smallest PM2.5 concentrations decrease.~~ Those isopleths present the same characteristics for ME and NI regions (~~Cf.~~ Supplementary information Figures S1, S4).

The qualitative evolution represented ~~on~~in Figure ~~3~~5 is further quantified by applying the ~~linear model~~GAM to the future meteorological variables in the regional climate projections. These results are represented by the probability density functions of the predicted concentrations of each GCM/RCM couple minus the estimated values for the historical ~~run~~simulation (e.g. 2071-2100 vs. 1976-2005, ~~Cf. Figure 4).~~Figure 6). For EA and ME, the longer tail of the probability density function of MPI-ESM-LR/CCLM ~~compare~~compared to the average of the models reflects that stronger pollution ~~episode~~episodes will occur in the future even if the mean of the concentrations ~~are~~is lower than the average of the ensemble ~~(Cf. Figure 4~~(Figure 6 for EA and Figure S2 for ME).

Besides the distribution, the ensemble mean and standard deviation of the estimated projected change in PM2.5 concentrations has been quantified (Table 3). All the selected regions depict a significant decrease of the PM2.5 concentrations across the multi-model proxy ensemble indicating that according to the GAM model, the climate benefit on particulate matter ~~in~~is a robust feature in these regions. The magnitude of the decrease depends on the region~~ and is expressed as follow~~, its ensemble mean (± standard deviation~~): -0.96 (± 0.18) µg/m³;~~) is -1.~~00~~08 (± 0.~~37~~21) µg/m³, -1.~~16~~03 (± 0.32) µg/m³, -0.83 ± (0.~~23~~14) µg/m³, for ~~resp.~~respectively EA, ME and NI ~~(Cf. ~~(Table 3).

In order to explain the differences in the response of individual RCM in the ensemble, we need to explore the historical meteorological variables probability density functions (PDF, Figure 7) and to compare them with the evolution of IPSL-CM5A-MR/WRF~~.~~ (Figure 5). The comparison of the historical distribution for the temperature reflects the stronger extremes of IPSL-CM5A-MR/WRF (e.g. colder than the others when it is cold). ~~Only in ~~It is only for the NI ~~our model~~region that IPSL-CM5A-MR/WRF lies in the mean of the ensemble. Concerning the PBL depth, the values are similar ~~than~~to the average of the ensemble for ME even if MPI-ESM-LR/RCA4 and EC-EARTH/RACMO2 present the largest values. IPSL-CM5A-MR/WRF has a thinner boundary layer for NI and a ~~larger for EA ~~deeper than the average for EA but the differences are limited ~~(Cf. ~~Figure ~~5~~7).

It is for CSIRO-Mk3-6-0/RCA4 ~~depicts~~that we find the most important decrease of PM2.5 for ~~all ~~the selected regions ~~except over NI where it is exceed by CanESM2/RCA4 (resp. 1.51 µg/m³ vs. 1.61 µg/m³; Cf. (~~Table 3). This is ~~linked~~related to a larger temperature rise ~~compare~~compared to the other models and a larger boundary layer height ~~evolution~~

~~compare~~increase compared to the other member of the ensemble for these regions ~~(Cf. ~~Figure ~~3~~5). CanESM2/RCA4 and CSIRO-Mk3-6-0/RCA4 ~~reflects~~exhibit the same ~~finding~~features for the ME region~~ ME~~.

MPI-ESM-LR/CCLM presents the smallest decrease of PM2.5 for each of the selected regions (e.g. over ME is almost 3 times smaller than the largest decrease) except EA where CNRM-CM5-LR/RCA4 presents a smaller decrease (-0.77 µg/m^3 vs. -0.81 µg/m^3). As already mentioned above, the particular tails of the statistically modelled PM2.5 distributions for EA and ME indicate ~~more~~a larger contribution of large pollution episodes in the future~~. The~~ for that RCM. But the historical distributions exhibit a larger boundary layer than the average models of the ensemble and a similar temperature. Thus, the low PM2.5 concentration decrease is explained by the ~~small~~limited average evolution of the meteorological drivers as shown ~~by the~~in Figure ~~3~~5~~. The evolution of the PBL depth depicts the relevance of this meteorological variable: a large part of the contour overlaps the red part of the background,. Hence it indicates more days with a thinner layer which is directly related to more PM25 pollution episodes.~~.

Overall we conclude that ~~the~~a climate benefit is ~~confirmed~~identified for the PM2.5 for each of the selected regions. ~~The~~To the extent that the statistical model is skillful, as demonstrated in Section 3.2.1, this result is robust across the range of available climate forcings since ~~all ~~the ~~proxy of~~whole ensemble ~~built with the bivariate statistical model applied to~~of regional climate projection present ~~similar density functions and average projected changes~~consistent features. The regional climate models that exhibit ~~a specific response~~the largest and smallest responses are CanESM2/RCA4; CSIRO-Mk3-6-0/RCA4 and MPI-ESM-LR/CCLM, which should therefore be considered in priority for ~~a more in-depth~~further evaluation using explicit deterministic projections involving full-frame regional climate and chemistry models.

4.2 Ozone peaks

For most of the selected regions (FR, IP, ME and NI,) the main drivers are the same (i.e. near surface temperature and short wave radiation~~) except for EA where the major drivers are~~). The isopleth in the background of Figure 5 show that temperature and ~~specific humidity. As discussed above for PM2.5, every figure (~~short wave radiation have a similar impact on ozone peaks, except in the larger range of short wave radiation anomalies, where temperature becomes less discriminating. All the isopleths (Figure 5 for EA and Figures S1, S4 ~~& Figure 3) shows~~and S7 for ME, NI, FR and IP) exhibit an ~~offset of the 2D density plot along the~~ increase in the distribution of temperatures ~~axe. The~~because the projected future is warmer than the historical period. According to the ozone peak concentrations predicted by the ~~linear model~~GAM (displayed in the background of ~~Figure 3)~~Figure 5) these ~~offsets~~increases will lead to more ozone episodes. This trend appears for the entire models ensemble so that we can conclude with confidence that ~~this~~the climate penalty bearing upon ozone is a robust feature even if the specific distribution ~~shape~~ of some of the models stand out (CanESM2/RCA4; CNRM-CM5-LR/RCA4; CSIRO-Mk3-6-0/RCA4; IPSL-CM5A-MR/WRF).

The ozone increase of the ensemble ~~is equal~~ reaches +10.~~11~~51 (± 3.~~22) μg/m^3, 8.23 (± 2.~~06) μg/m^3, ~~9.23~~+11.70 (± 3.63) μg/m^3, +11.53 (± 1.~~13~~55) μg/m^3, ~~6~~+9.86 (± 4.41 ~~(± 2.14~~) μg/m^3, ~~7.43 (± 2.02~~+4.82 (± 1.79) μg/m^3 for EA, FR, IP, ME and NI ~~(Cf.~~ (Table 3). These values confirm the statistically significant climate penalty (the mean is at least two times larger than the standard deviation). However, as already mentioned for Figure ~~3~~5, we find minor differences among the models. ~~Here the meteorological variables and their evolution are discussed to explain these differences.~~ The meteorological distributions are ~~slightly~~marginally different between the models of the ensemble: the summertime temperature predicted by IPSL-CM5A-MR/WRF has stronger extremes than the other models. Moreover, it is warmer than the ensemble in EA. ~~The specific humidity is around 1.5 times larger for IPSL-CM5A-MR/WRF than for the other models.~~ Concerning ~~the last meteorological variable,~~ incoming short wave radiation, IPSL-CM5A-MR/WRF lies in the average (~~Cf.~~ Figure S3, S6, S9~~). Only~~) except for the region EA where the amount of incoming radiation is the highest among the ensemble (Figure 7). Note that, only EC-EARTH/RACMO2 and MPI-ESM-LR/RCA4 exhibits lower values (around half of the average for MPI-ESM-LR/CCLM). The lower amount of summertime incoming short wave radiation for the couple MPI-ESM-LR/CCLM is relevant for all the selected regions.

The magnitude of the ozone rise ~~changes~~differs between the models and the regions. Note that CanESM2/RCA4 exhibits the ~~most important discrepancy~~largest difference (i.e. around 1.5 times the ensemble mean) followed by CSIRO-Mk3-6-0/RCA4 for each selected regions. This is explained by the ~~significant~~larger temperature increase during summertime which is the major driver, as identified by the statistical models, of ozone concentration. Note that the value is skyrocketing for the region ~~EA when specific humidity is the 2nd predictor, 4~~ME, 5 times the value of IPSL-CM5A-MR/WRF~~,~~ which shows one of the ~~lower increase~~lowest increases. CNRM-CM5-LR/RCA4 presents the ~~2nd~~ lowest increase.

On the contrary, the lower increase of the summer temperature and sometimes a decrease of the incoming short wave radiation amount (e.g. IPSL-CM5A-MR/WRF in NI) ~~are associated~~lead to lower ozone concentration changes for IPSL-CM5A-MR/WRF and CNRM-CM5-LR/RCA4 for FR, IP, ME and NI ~~(Cf.~~ (Table 3~~).~~). Note the specific evolution for the region NI, where the IPSL-CM5A-MR/WRF model yields almost no increase of the ozone concentration compared to the other models while on the map of the differences in the deterministic model (Figure 1.f), the evolution was statistically significant. This absence of evolution reflects the limitation of the statistical models.

~~On the~~In figure S5, we can point out ~~the particular~~an outstanding pattern of the MPI-ESM-LR/CCLM distribution for the NI region~~: a wide and flat Gaussian~~ with particularly large tails. The ozone rise would be more pronounced for the upper quantile which depicts more extreme ~~polluted episode.~~ozone pollution episode (note that this was also the case for that model in terms of PM2.5 pollution).

Overall the climate penalty is confirmed even if some regional climate models stand out of the distribution, such as CanESM2/RCA4; CNRM-CM5-LR/RCA4 and CSIRO-Mk3-6-0/RCA4 which should therefore be considered for further deterministic projections.

5 Conclusions

An alternative technique to assess the robustness of projections of the impact of climate change on air quality has been introduced. Using a training dataset consisting of long-term deterministic regional climate and air quality projections, we could build ~~simple~~ statistical models of the response of ozone and particulate pollution to the main ~~climate~~meteorological drivers for several regions of Europe. Applying such statistical models to an ensemble of regional climate ~~projection~~projections leads to the development of an ensemble of proxy projections of air quality under various future climate ~~forcing~~forcings. The assessment of the spread of the ensemble of proxy projections allows inferring the robustness of the impact of climate change, as well as selecting a subset of climate models to be used in priority for future explicit air quality projections, therefore proposing ~~a smart~~an optimized exploration of the ensemble.

The main ~~climate~~meteorological drivers that were identified are (i) for PM2.5: the boundary layer depth and the near surface temperature and (ii) for ozone: the near surface temperature and the incoming short wave radiation~~ except for Eastern Europe where specific humidity is the second predictor.~~. The skill of the statistical models depends on the regions of Europe and the pollutant.

For PM2.5 and the regions Eastern Europe (EA) and Mid Europe (ME), a ~~bivariate linear least square~~generalized additive model captures about ~~50~~60% of the variance~~.~~ and for Northern Italy 40%. But for British Isles (BI) and Scandinavia (SC), where air pollution is largely driven by long range transport, such a ~~simple and~~ local approach is not able to reproduce the variability of pollutant concentrations.

The ozone concentrations are well reproduced by the statistical model for the following regions: Eastern Europe (EA), France (FR), Iberian Peninsula (IP), Mid Europe (ME) and ~~Northen~~Northern Italy (NI). The meteorological variables are not discriminating enough ~~to depict~~for the ~~pollutant concentration for~~ Mediterranean region. For the regions where the performances of the statistical model were considered satisfactory, a proxy of the future pollutant concentrations could be estimated (i.e. (i) EA, ME and NI (ii) EA, FR, IP, ME and NI).

An overall climate benefit for PM2.5 was found in the proxy ensemble of climate and air quality projections. The ensemble mean change is~~ -0.96 (standard deviation: ± 0.18 µg/m³),~~ -1.~~00~~08 (± 0.~~37~~21) µg/m³, -1.~~16 (±~~03 (± 0.32) µg/m³, -0.83 ± (0.~~23~~14) µg/m³, for ~~resp.~~respectively EA, ME and NI. This beneficial impact of climate change for particulate matter pollution is in agreement with the deterministic projections of (Huszar et al., 2011;Juda-Rezler et al., 2012;Colette et al., 2013) but in opposition to (Manders et al., 2012). These differences could be partly explained by the different time windows (i.e. 2060 -2041 vs. 2100-2071), climate scenario (i.e. A1B which is similar to RCP6.0 vs. RCP8.5) and pollutant (i.e. PM10 vs. PM2.5). This impact of climate change on particulate pollution should be put in perspective with the magnitude of the change that is expected from the current air quality legislation. Such a comparison was performed by (Colette et al., 2013) who found (on average over Europe) a climate benefit by the middle of the century of the order of

0-1 µg/m3, therefore in line with our estimate but also much lower than the expected reduction of 7-8µg/m3 that they attributed to air quality policies.

For all the selected regions a robust climate penalty on ozone was identified: +10.~~11~~51 (± 3.~~22) µg/m^3, 8.23 (± 2.~~06) µg/m^3, ~~9.23~~+11.70 (± 3.63) µg/m^3, +11.53 (± 1.~~13~~55) µg/m^3, ~~6~~+9.86 (± 4.41~~ (± 2.14~~) µg/m^3, ~~7.43 (± 2.02~~+4.82 (± 1.79) µg/m^3 for ~~resp.~~respectively EA, FR, IP, ME and NI. This finding is in line with previous studies (Meleux et al., 2007;Huszar et al., 2011;Katragkou et al., 2011;Jiménez-Guerrero et al., 2012;Juda-Rezler et al., 2012;Langner et al., 2012a;Langner et al., 2012b;Colette et al., 2013;Hedegaard et al., 2013;Varotsos et al., 2013;Colette et al., 2015)~~.~~

~~The~~. It should be noted that when comparing the impact of climate change and emission reduction strategies, (Colette et al., 2013) found a climate penalty of the order of 2-3µg/m3 (which is broadly consistent with our results given that they focused on the middle of the century) that could be compensated with the expected magnitude of the reduction of 5-10µg/m3 brought about by air quality policies.

The major strength of our approach is to account for the climate uncertainty in the recent EuroCordex ensemble of regional climate projections, whereas all the published literature relied on very limited subset of RCM forcing (at best two for a given chemistry-transport modelling study). We therefore propose an unprecedented view in the robustness of the impact of climate change on air quality ~~inferred from~~ across an ensemble of climate forcing. However, this ~~proxy of ensemble cannot be considered as a very definitive statement given that~~ achievement is limited by the quality of the underlying statistical model that does not capture all the variance of the air quality response to climate change. ~~The somewhat simple structure of the statistical model and the use of a single set of~~ These results should thus be ultimately compared with further deterministic ~~projection for its training/validation, are additional limitations of the approach. However, besides this~~projections using a range of climate forcings. Then, our approach can yield precious information ~~on the robustness, this proxy approach also allows~~in pointing out which regional climate models ~~that~~ should be investigated in priority~~ in the context of deterministic model projection~~, therefore proposing a smart exploration of the ensemble of projections. The following models: CanESM2/RCA4; CNRM-CM5-LR/RCA4 and CSIRO-Mk3-6-0/RCA4 and MPI-ESM-LR/CCLM, have been identified as the climate models that should be used in priority for future air quality.

Finally, ~~this~~we should add that the method~~,~~ applied here for air quality projection also opens ~~also~~ the way for ~~such approaches in~~ other climate impact studies, where quantifying uncertainties using low computational demand is desirable.

Acknowledgements

The climate and air quality projections used in the present study were produced as part of the European Union's Seventh Framework Programme (FP7/2007-2013) under grant agreement no. 282687 (ATOPICA) and computing resources of the TGCC/CCRT/CEA.

References

The references highlighted in yellow have been added to the revised manuscript.

Bessagnet, B., Seigneur, C., and Menut, L.: Impact of dry deposition of semi-volatile organic compounds on secondary organic aerosols, Atmospheric Environment, 44, 1781-1787, 10.1016/j.atmosenv.2010.01.027, 2010.

Bessagnet, B., A. Colette, F. Meleux, L. Rouïl, A. Ung, O. Favez, C. Cuvelier, P. Thunis, S. Tsyro, R. Stern, A. Manders, R. Kranenburg, A. Aulinger, J. Bieser, M. Mircea, G. Briganti, A. Cappelletti, G. Calori, S. Finardi, C. Silibello, G. Ciarelli, S. Aksoyoglu, A. Prévôt, M.-T. Pay, J. M. Baldasano, M. García Vivanco, J. L. Garrido, I. Palomino, F. Martín, G. Pirovano, P. Roberts, L. Gonzalez, L. White, L. Menut, J.-C. Dupont, C. Carnevale, and Pederzoli, A.: The EURODELTA III exercise - Model evaluation with observations issued from the 2009 EMEP intensive period and standard measurements in Feb/Mar 2009, 2014.

Cattiaux, J., Douville, H., and Peings, Y.: European temperatures in CMIP5: origins of present-day biases and future uncertainties, Clim Dyn, 41, 2889-2907, 10.1007/s00382-013-1731-y, 2013.

Christensen, J., and Christensen, O.: A summary of the PRUDENCE model projections of changes in European climate by the end of this century, Climatic Change, 81, 7-30, 10.1007/s10584-006-9210-7, 2007.

Colette, A., Granier, C., Hodnebrog, Ø., Jakobs, H., Maurizi, A., Nyiri, A., Bessagnet, B., D'Angiola, A., D'Isidoro, M., Gauss, M., Meleux, F., Memmesheimer, M., Mieville, A., Rouïl, L., Russo, F., Solberg, S., Stordal, F., and Tampieri, F.: Air quality trends in Europe over the past decade: a first multi-model assessment, Atmos. Chem. Phys., 11, 11657-11678, 10.5194/acp-11-11657-2011, 2011.

Colette, A., Bessagnet, B., Vautard, R., Szopa, S., Rao, S., Schucht, S., Klimont, Z., Menut, L., Clain, G., Meleux, F., Curci, G., and Rouïl, L.: European atmosphere in 2050, a regional air quality and climate perspective under CMIP5 scenarios, Atmos. Chem. Phys., 13, 7451-7471, 10.5194/acp-13-7451-2013, 2013.

Colette, A., Andersson, C., Baklanov, A., Bessagnet, B., Brandt, J., H. Christensen, J., Doherty, R., Engardt, M., Geels, C., Giannakopoulos, C., B. Hedegaard, G., Katragkou, E., Langner, J., Lei, H., Manders, A., Melas, D., Meleux, F., Rouil, L., Sofiev, M., Soares, J., S. Stevenson, D., Tombrou-Tzella, M., V. Varotsos, K., and Young, P.: Is the ozone climate penalty robust in Europe?, Environmental Research Letters, 10, 084015, 10.1088/1748-9326/10/8/084015, 2015.

Cuvelier, C., Thunis, P., Vautard, R., Amann, M., Bessagnet, B., Bedogni, M., Berkowicz, R., Brandt, J., Brocheton, F., Builtjes, P., Carnavale, C., Coppalle, A., Denby, B., Douros, J., Graf, A., Hellmuth, O., Hodzic, A., Honoré, C., Jonson, J., Kerschbaumer, A., de Leeuw, F., Minguzzi, E., Moussiopoulos, N., Pertot, C., Peuch, V. H., Pirovano, G., Rouil, L., Sauter, F., Schaap, M., Stern, R., Tarrason, L., Vignati, E., Volta, M., White, L., Wind, P., and Zuber, A.: CityDelta: A model intercomparison study to explore the impact of emission reductions in European cities in 2010, Atmospheric Environment, 41, 189-207, 2007.

Dockery, D. W., and Pope, C. A.: Acute Respiratory Effects of Particulate Air Pollution, Annual Review of Public Health, 15, 107-132, doi:10.1146/annurev.pu.15.050194.000543, 1994.

Doherty, R. M., Wild, O., Shindell, D. T., Zeng, G., MacKenzie, I. A., Collins, W. J., Fiore, A. M., Stevenson, D. S., Dentener, F. J., Schultz, M. G., Hess, P., Derwent, R. G., and Keating, T. J.: Impacts of climate change on surface ozone and intercontinental ozone pollution: A multi-model study, Journal of Geophysical Research: Atmospheres, 118, 3744-3763, 10.1002/jgrd.50266, 2013.

Dufresne, J. L., Foujols, M. A., Denvil, S., Caubel, A., Marti, O., Aumont, O., Balkanski, Y., Bekki, S., Bellenger, H., Benshila, R., Bony, S., Bopp, L., Braconnot, P., Brockmann, P., Cadule, P., Cheruy, F., Codron, F., Cozic, A., Cugnet, D., de Noblet, N., Duvel, J. P., Ethé, C., Fairhead, L., Fichefet, T., Flavoni, S., Friedlingstein, P., Grandpeix, J. Y., Guez, L., Guilyardi, E., Hauglustaine, D., Hourdin, F., Idelkadi, A., Ghattas, J., Joussaume, S., Kageyama, M., Krinner, G., Labetoulle, S., Lahellec, A., Lefebvre, M. P., Lefevre, F., Levy, C., Li, Z. X., Lloyd, J., Lott, F., Madec, G., Mancip, M., Marchand, M., Masson, S., Meurdesoif, Y., Mignot, J., Musat, I., Parouty, S., Polcher, J., Rio, C., Schulz, M., Swingedouw, D., Szopa, S., Talandier, C., Terray, P., Viovy, N., and Vuichard, N.: Climate change projections using the IPSL-CM5 Earth System Model: from CMIP3 to CMIP5, Clim Dyn, 40, 2123-2165, 10.1007/s00382-012-1636-1, 2013.

Fountoukis, C., and Nenes, A.: ISORROPIA II: a computationally efficient thermodynamic equilibrium model for K+ Ca2+ Mg2+ NH4+ Na+ SO42- NO3- Cl- H2O aerosols, Atmos. Chem. Phys., 7, 4639-4659, 10.5194/acp-7-4639-2007, 2007.

Hauglustaine, D. A., Balkanski, Y., and Schulz, M.: A global model simulation of present and future nitrate aerosols and their direct radiative forcing of climate, Atmos. Chem. Phys., 2014.

Hedegaard, G. B., Christensen, J. H., and Brandt, J.: The relative importance of impacts from climate change vs. emissions change on air pollution levels in the 21st century, Atmos. Chem. Phys., 13, 3569-3585, 2013.

Huszar, P., Juda-Rezler, K., Halenka, T., Chervenkov, H., Syrakov, D., Krüger, B., Zanis, P., Melas, D., Katragkou, E., Reizer, M., Trapp, W., and Belda, M.: Effects of climate change on ozone and particulate matter over Central and Eastern Europe, Climate Research, 50, 51-68, 10.3354/cr01036, 2011.

Jacob, D., Petersen, J., Eggert, B., Alias, A., Christensen, O., Bouwer, L., Braun, A., Colette, A., Déqué, M., Georgievski, G., Georgopoulou, E., Gobiet, A., Menut, L., Nikulin, G., Haensler, A., Hempelmann, N., Jones, C., Keuler, K., Kovats, S., Kröner, N., Kotlarski, S., Kriegsmann, A., Martin, E., van Meijgaard, E., Moseley, C., Pfeifer, S., Preuschmann, S., Radermacher, C., Radtke, K., Rechid, D., Rounsevell, M., Samuelsson, P., Somot, S., Soussana, J.-F., Teichmann, C., Valentini, R., Vautard, R., Weber, B., and Yiou, P.: EURO-CORDEX: new high-resolution climate change projections for European impact research, Regional Environmental Change, 14, 563-578, 10.1007/s10113-013-0499-2, 2014.

Jacob, D. J., and Winner, D. A.: Effect of climate change on air quality, Atmospheric Environment, 43, 51-63, http://dx.doi.org/10.1016/j.atmosenv.2008.09.051, 2009.

Jerrett, M., Burnett, R. T., Pope, C. A., Ito, K., Thurston, G., Krewski, D., Shi, Y., Calle, E., and Thun, M.: Long-Term Ozone Exposure and Mortality, New England Journal of Medicine, 360, 1085-1095, doi:10.1056/NEJMoa0803894, 2009.

Jiménez-Guerrero, P., Montávez, J. P., Gómez-Navarro, J. J., Jerez, S., and Lorente-Plazas, R.: Impacts of climate change on ground level gas-phase pollutants and aerosols in the Iberian Peninsula for the late XXI century, Atmospheric Environment, 55, 483-495, http://dx.doi.org/10.1016/j.atmosenv.2012.02.048, 2012.

Juda-Rezler, K., Reizer, M., Huszar, P., Krüger, B. C., Zanis, P., Syrakov, D., Katragkou, E., Trapp, W., Melas, D., and Chervenkov, H.: Modelling the effects of climate change on air quality over Central and Eastern Europe: concept, evaluation and projections, Climate Research, 53, 179-203, 2012.

Jury, M.: Evaluation of Global Climate Models as Regional Climate Model Drivers, Wegener Center for Climate and Global Change, University of Graz, 2012.

Katragkou, E., Zanis, P., Kioutsioukis, I., Tegoulias, I., Melas, D., Krüger, B. C., and Coppola, E.: Future climate change impacts on summer surface ozone from regional climate-air quality simulations over Europe, Journal of Geophysical Research: Atmospheres, 116, D22307, 10.1029/2011JD015899, 2011.

Katragkou, E., García-Díez, M., Vautard, R., Sobolowski, S., Zanis, P., Alexandri, G., Cardoso, R. M., Colette, A., Fernandez, J., Gobiet, A., Goergen, K., Karacostas, T., Knist, S., Mayer, S., Soares, P. M. M., Pytharoulis, I., Tegoulias, I., Tsikerdekis, A., and Jacob, D.: Regional climate hindcast simulations within EURO-CORDEX: evaluation of a WRF multi-physics ensemble, Geosci. Model Dev., 8, 603-618, 10.5194/gmd-8-603-2015, 2015.

Klimont, Z., Kupiainen, K., Heyes, C., Cofala, J., Rafaj, P., Höglund-Isaksson, L., Borken, J., Schöpp, W., Winiwarter, W., Purohit, P., Bertok, I., and Sander R.: ECLIPSE V4a: Global emission data set developed with the GAINS model for the period 2005 to 2050 Key features and principal data sources International Institute for Applied Systems Analysis (IIASA), Schlossplatz, 1, 2361, 2013.

Kotlarski, S., Keuler, K., Christensen, O. B., Colette, A., Déqué, M., Gobiet, A., Goergen, K., Jacob, D., Lüthi, D., van Meijgaard, E., Nikulin, G., Schär, C., Teichmann, C., Vautard, R., Warrach-Sagi, K., and Wulfmeyer, V.: Regional climate modeling on European scales: a joint standard evaluation of the EURO-CORDEX RCM ensemble, Geosci. Model Dev., 7, 1297-1333, 10.5194/gmd-7-1297-2014, 2014.

Lacressonniere, G., Foret, G., Beekmann, M., Siour, G., Engardt, M., Gauss, M., Watson, L., Andersson, C., Colette, A., Josse, B., Marécal, V., Nyiri, A., and Vautard, R.: Impacts of regional climate change on air quality projections and associated uncertainties Climatic Change, Accepted, 2016.

Langner, J., Engardt, M., and Andersson, C.: European summer surface ozone 1990-2100, Atmos. Chem. Phys., 12, 10097-10105, 2012a.

Langner, J., Engardt, M., Baklanov, A., Christensen, J. H., Gauss, M., Geels, C., Hedegaard, G. B., Nuterman, R., Simpson, D., Soares, J., Sofiev, M., Wind, P., and Zakey, A.: A multi-model study of impacts of climate change on surface ozone in Europe, Atmos. Chem. Phys., 12, 10423-10440, 10.5194/acp-12-10423-2012, 2012b.

Lecœur, È., and Seigneur, C.: Dynamic evaluation of a multi-year model simulation of particulate matter concentrations over Europe, Atmos. Chem. Phys., 2013.

Lecœur, È., Seigneur, C., Pagé, C., and Terray, L.: A statistical method to estimate PM2.5 concentrations from meteorology and its application to the effect of climate change, Journal of Geophysical Research: Atmospheres, 119, 3537-3585, 10.1002/2013JD021172, 2014.

Manders, A. M. M., van Meijgaard, E., Mues, A. C., Kranenburg, R., van Ulft, L. H., and Schaap, M.: The impact of differences in large-scale circulation output from climate models on the regional modeling of ozone and PM, Atmos. Chem. Phys., 12, 9441-9458, 10.5194/acp-12-9441-2012, 2012.

McGrath-Spangler, E. L., Molod, A., Ott, L. E., and Pawson, S.: Impact of planetary boundary layer turbulence on model climate and tracer transport, Atmos. Chem. Phys., 15, 7269-7286, 10.5194/acp-15-7269-2015, 2015.

Megaritis, A. G., Fountoukis, C., Charalampidis, P. E., Denier van der Gon, H. A. C., Pilinis, C., and Pandis, S. N.: Linking climate and air quality over Europe: effects of meteorology on PM2.5 concentrations, Atmos. Chem. Phys. Discuss., 14, 10345-10391, 10.5194/acpd-14-10345-2014, 2014.

Meleux, F., Solmon, F., and Giorgi, F.: Increase in summer European ozone amounts due to climate change, Atmospheric Environment, 41, 7577-7587, 2007.

Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J. L., Pison, I., Siour, G., Turquety, S., Valari, M., Vautard, R., and Vivanco, M. G.: CHIMERE 2013: a model for regional atmospheric composition modelling, Geosci. Model Dev., 6, 981-1028, 10.5194/gmd-6-981-2013, 2013.

Prybutok, V. R., Yi, J., and Mitchell, D.: Comparison of neural network models with ARIMA and regression models for prediction of Houston's daily maximum ozone concentrations, European Journal of Operational Research, 122, 31-40, http://dx.doi.org/10.1016/S0377-2217(99)00069-7, 2000.

Rouïl, L., Honore, C., Vautard, R., Beekmann, M., Bessagnet, B., Malherbe, L., Meleux, F., Dufour, A., Elichegaray, C., Flaud, J. M., Menut, L., Martin, D., Peuch, A., Peuch, V. H., and Poisson, N.: PREV'AIR An Operational Forecasting and Mapping System for Air Quality in Europe, Bulletin of the American Meteorological Society, 90, 73-83, 10.1175/2008bams2390.1, 2009.

Schlink, U., Herbarth, O., Richter, M., Dorling, S., Nunnari, G., Cawley, G., and Pelikan, E.: Statistical models to assess the health effects and to forecast ground-level ozone, Environmental Modelling & Software, 21, 547-558, http://dx.doi.org/10.1016/j.envsoft.2004.12.002, 2006.

Seinfeld, J. H., and Pandis, S. N.: Atmospheric Chemistry and Physics : From Air Pollution to Climate Change, 2008.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X. Y., Wang, W., and Powers, J. G.: A Description of the Advanced Research WRF Version 3, NCAR, Boulder, Colorado, USA, 2008.

Slini, T., Kaprara, A., Karatzas, K., and Moussiopoulos, N.: PM10 forecasting for Thessaloniki, Greece, Environmental Modelling & Software, 21, 559-565, http://dx.doi.org/10.1016/j.envsoft.2004.06.011, 2006.

Solazzo, E., Bianconi, R., Pirovano, G., Matthias, V., Vautard, R., Moran, M. D., Wyat Appel, K., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Miranda, A. I., Nopmongcol, U., Prank, M., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Operational model evaluation for particulate matter in Europe and North America in the context of AQMEII, Atmospheric Environment, 53, 75-92, 2012a.

Solazzo, E., Bianconi, R., Vautard, R., Appel, K. W., Moran, M. D., Hogrefe, C., Bessagnet, B., Brandt, J. r., Christensen, J. H., Chemel, C., Coll, I., Denier van der Gon, H., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Jericevic, A., Kraljevic, L., Miranda, A. I., Nopmongcol, U., Pirovano, G., Prank, M., Riccio, A., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S. T., and Galmarini, S.: Model evaluation and ensemble modelling of surface-level ozone in Europe and North America in the context of AQMEII, Atmospheric Environment, 53, 60-74, 2012b.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A.: An Overview of CMIP5 and the Experiment Design, Bulletin of the American Meteorological Society, 93, 485-498, 2012.

Thunis, P., Pederzoli, A., and Pernigotti, D.: Performance criteria to evaluate air quality modeling applications, Atmospheric Environment, 59, 476-482, http://dx.doi.org/10.1016/j.atmosenv.2012.05.043, 2012.

Thunis, P., Clappier, A., Pisoni, E., and Degraeuwe, B.: Quantification of non-linearities as a function of time averaging in regional air quality modeling applications, Atmospheric Environment, 103, 263-275, http://dx.doi.org/10.1016/j.atmosenv.2014.12.057, 2015.

van Loon, M., Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P. J. H., Christensen, J. H., Cuvelier, C., Graff, A., Jonson, J. E., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrasón, L., Thunis, P., Vignati, E., White, L., and Wind, P.: Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble, Atmospheric Environment, 41, 2083-2097, http://dx.doi.org/10.1016/j.atmosenv.2006.10.073, 2007.

van Vuuren, D., Edmonds, J., Kainuma, M., Riahi, K., Thomson, A., Hibbard, K., Hurtt, G., Kram, T., Krey, V., Lamarque, J.-F., Masui, T., Meinshausen, M., Nakicenovic, N., Smith, S., and Rose, S.: The representative concentration pathways: an overview, Climatic Change, 109, 5-31, 10.1007/s10584-011-0148-z, 2011.

Vano, J., and Lettenmaier, D.: A sensitivity-based approach to evaluating future changes in Colorado River discharge, Climatic Change, 122, 621-634, 10.1007/s10584-013-1023-x, 2014.

Varotsos, K. V., Giannakopoulos, C., and Tombrou, M.: Assessment of the Impacts of Climate Change on European Ozone Levels, Water, Air, & Soil Pollution C7 - 1596, 224, 1-13, 2013.

Vautard, R., Honoré, C., Beekmann, M., and Rouil, L.: Simulation of ozone during the August 2003 heat wave and emission control scenarios, Atmospheric Environment, 39, 2957-2967, 2005.

Vautard, R., Maidi, M., Menut, L., Beekmann, M., and Colette, A.: Boundary layer photochemistry simulated with a two-stream convection scheme, Atmospheric Environment, 41, 8275-8287, 2007.

Watson, L., Lacressonnière, G., Gauss, M., Engardt, M., Andersson, C., Josse, B., Marécal, V., Nyiri, A., Sobolowski, S., Siour, G., and Vautard, R.: The impact of meteorological forcings on gas phase air pollutants over Europe, Atmospheric Environment, 119, 240-257, http://dx.doi.org/10.1016/j.atmosenv.2015.07.037, 2015.

Wilson, R. C., Fleming, Z. L., Monks, P. S., Clain, G., Henne, S., Konovalov, I. B., Szopa, S., and Menut, L.: Have primary emission reduction measures reduced ozone across Europe? An analysis of European rural background ozone trends 1996-2005, Atmos. Chem. Phys., 12, 437-454, 2012.

Wood, S. N.: Generalized additive models: an introduction with R, 2006.

Young, P. J., Archibald, A. T., Bowman, K. W., Lamarque, J. F., Naik, V., Stevenson, D. S., Tilmes, S., Voulgarakis, A., Wild, O., Bergmann, D., Cameron-Smith, P., Cionni, I., Collins, W. J., Dalsoren, S. B., Doherty, R. M., Eyring, V., Faluvegi, G., Horowitz, L. W., Josse, B., Lee, Y. H., MacKenzie, I. A., Nagashima, T., Plummer, D. A., Righi, M., Rumbold, S. T., Skeie, R. B., Shindell, D. T., Strode, S. A., Sudo, K., Szopa, S., and Zeng, G.: Pre-industrial to end 21st century projections of tropospheric ozone from the Atmospheric Chemistry and Climate Model Intercomparison Project (ACCMIP), Atmos. Chem. Phys. Discuss., 12, 21615-21677, 2012.

All the following tables and figures have changed except figure 7.

PM25			
Regions	R^2	Meteorological variable 1	Meteorological variable 2
BI	0,327	PBL-height	Surface wind
IP	0,228	PBL-height	Specific humidity
FR	0,343	PBL-height	Near surface temperature
ME	0,613	PBL-height	Near surface temperature
SC	0,206	Specific humidity	Incoming short wave radiation
NI	0,403	PBL-height	Near surface temperature
MD	0,194	PBL-height	Surface wind
EA	0,595	PBL-height	Near surface temperature

Table 1: Statistical models per region that explain the average PM2.5 concentrations during 1976-2005.

Ozone max			
Regions	R^2	Meteorological variable 1	Meteorological variable 2
BI	0,402	Incoming short wave radiation	Specific humidity
IP	0,543	Near surface temperature	Incoming short wave radiation
FR	0,579	Near surface temperature	Incoming short wave radiation
ME	0,709	Near surface temperature	Incoming short wave radiation
SC	0,228	Near surface temperature	PBL-height
NI	0,603	Incoming short wave radiation	Near surface temperature
MD	0,343	Near surface temperature	Surface wind
EA	0,671	Near surface temperature	Incoming short wave radiation

Table 2: Statistical models per region that explain the daily maximum summer ozone levels during 1976-2005.

RCP8.5 2071-2100	Delta (future - historical)							
	Ozone max					PM2.5		
GCM/RCM \ Regions	EA	FR	IP	ME	NI	EA	ME	NI
CNRM-CM5-LR/RCA4	8,00	6,96	9,75	4,82	3,69	-0,77	-0,82	-0,71
CSIRO-Mk3-6-0/RCA4	11,26	16,03	13,30	14,15	5,81	-1,39	-1,72	-1,06
CanESM2/RCA4	17,97	19,03	15,07	21,20	7,46	-1,29	-1,56	-1,03
EC-EARTH/RACMO2	7,77	11,37	10,79	8,55	6,77	-1,16	-0,98	-0,77
EC-EARTH/RCA4	10,88	14,43	11,45	12,11	5,15	-0,92	-0,92	-0,75
GFDL-ESM2M/RCA4	7,26	7,79	10,28	5,85	4,54	-1,04	-0,90	-0,70
IPSL-CM5A-MR/RCA4	13,76	13,46	12,88	11,02	4,43	-1,28	-1,12	-1,04
IPSL-CM5A-MR/WRF	10,11	6,05	9,08	5,19	0,01	-1,32	-1,30	-0,86
MIROC5/RCA4	12,30	11,29	11,61	9,62	3,85	-1,16	-0,86	-0,85
MPI-ESM-LR/CCLM	6,40	9,63	11,03	6,01	5,58	-0,81	-0,58	-0,62
MPI-ESM-LR/RCA4	9,56	11,75	11,51	9,64	5,54	-1,02	-0,79	-0,83
NorESM1-M/RCA4	10,88	12,60	11,58	10,12	5,02	-0,79	-0,88	-0,76
Ensemble Mean	10,51	11,70	11,53	9,86	4,82	-1,08	-1,03	-0,83
Ensemble Standard Deviation	3,06	3,63	1,55	4,41	1,79	0,21	0,32	0,14

Table 3: Predicted concentrations evolution of summertime ozone and PM2.5 (expressed in μg/m^3) per selected regions and per model. The ensemble mean and standard deviation are also calculated.

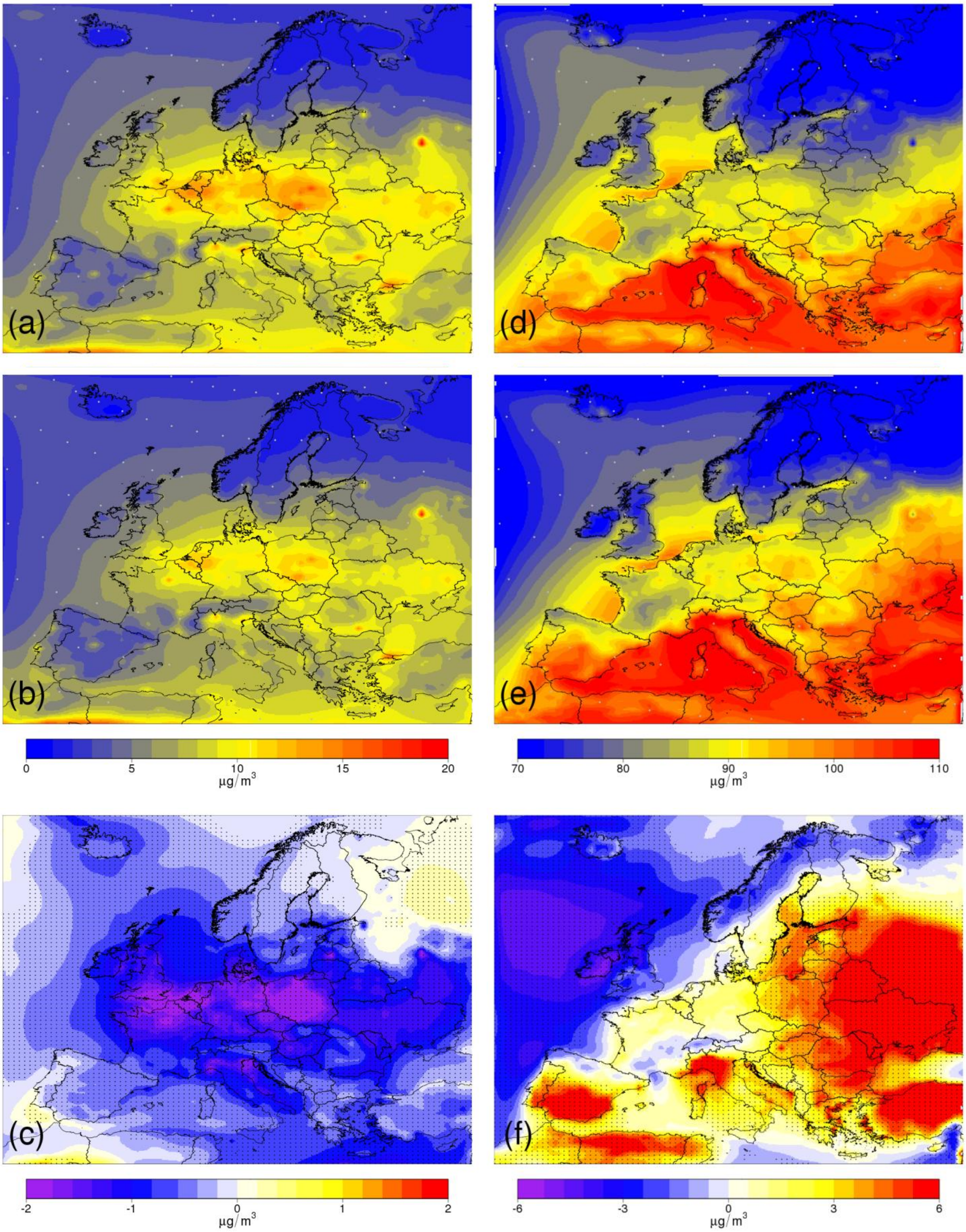

Figure 1: The left column represents daily average PM2.5 concentrations for the historical (1976-2005) (a), the end of the century (RCP8.5 - 2071-2100) (b) and the difference between the future and the historical (c). The statistical significance of this difference is evaluated by a t-test and represented by a

black point. The right column presents the same figure for daily maximum ozone projections. For both pollutants, the CTM CHIMERE has been used to predict the concentration (Section 2.2).

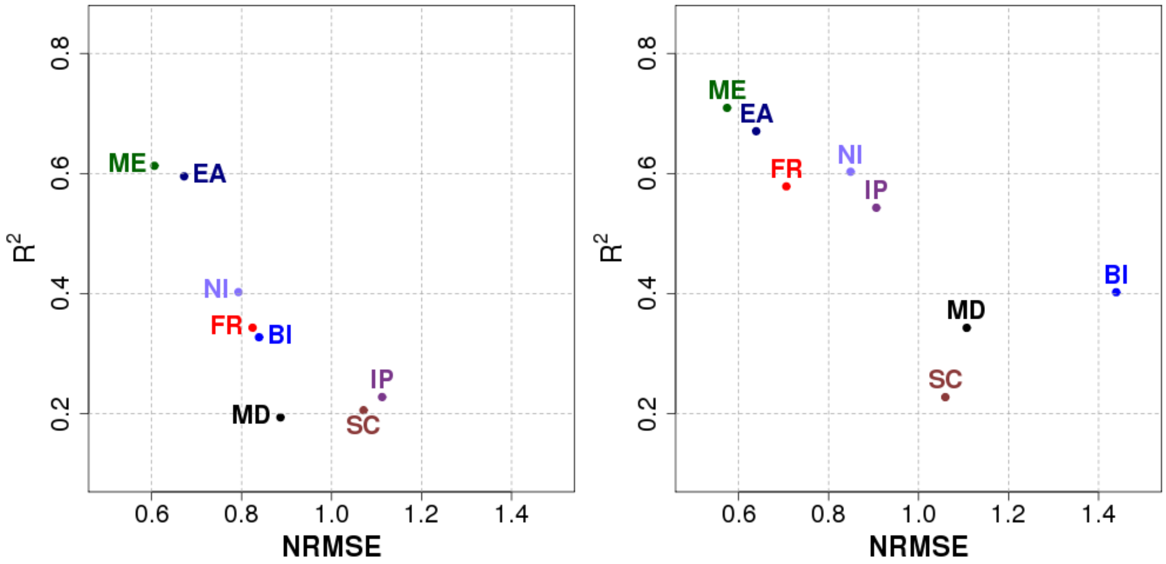

Figure 2: Statistical model evaluation for PM2.5 (left) and ozone (right). The x-axis represents the Normalized Mean Square Error applied to the delta (future minus historical) of the generalized additive model and CHIMERE. The y-axis represents the R^2 of the statistical model (training period).

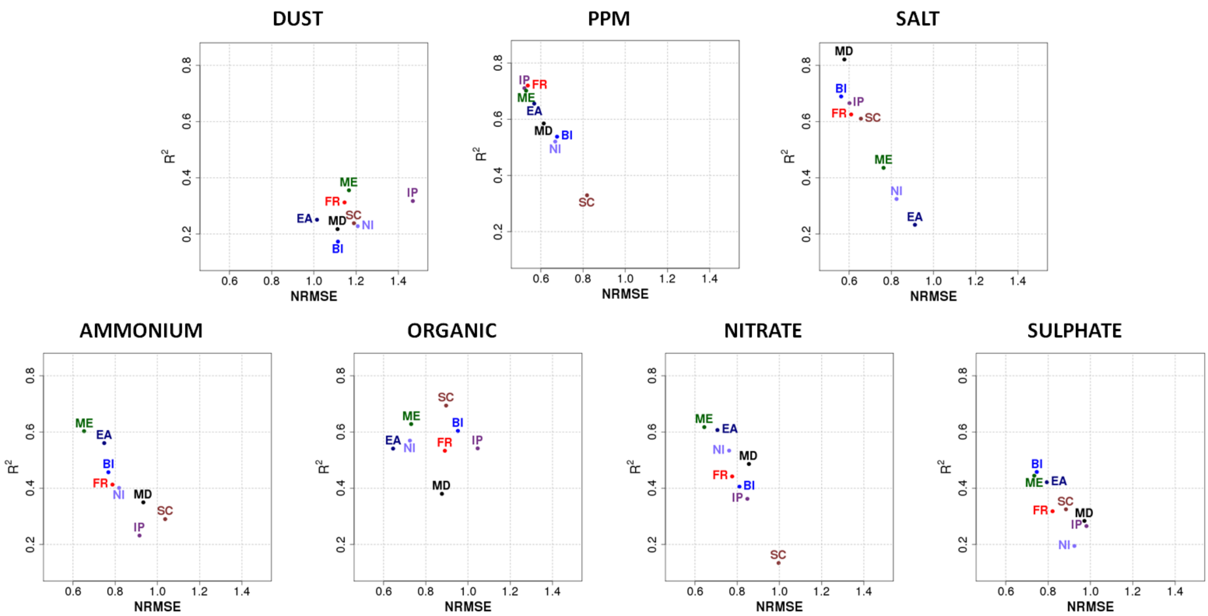

Figure 3: Statistical model evaluation for each particulate matter constituent (from left to right: Dust, Primary Particulate Matter, Sea-salt, Ammonium, Organic fraction, Nitrate, Sulphate). The x-axis represents the Normalized Mean Square Error applied to the delta (future minus historical) of either the

generalized additive model or CHIMERE. The y-axis represents the R^2 of the statistical model (training period).

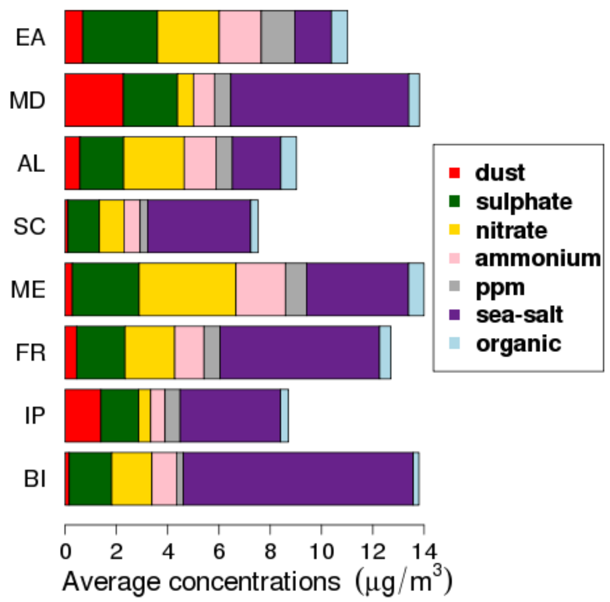

Figure 4: Average particulate matter composition for the historical period per region.

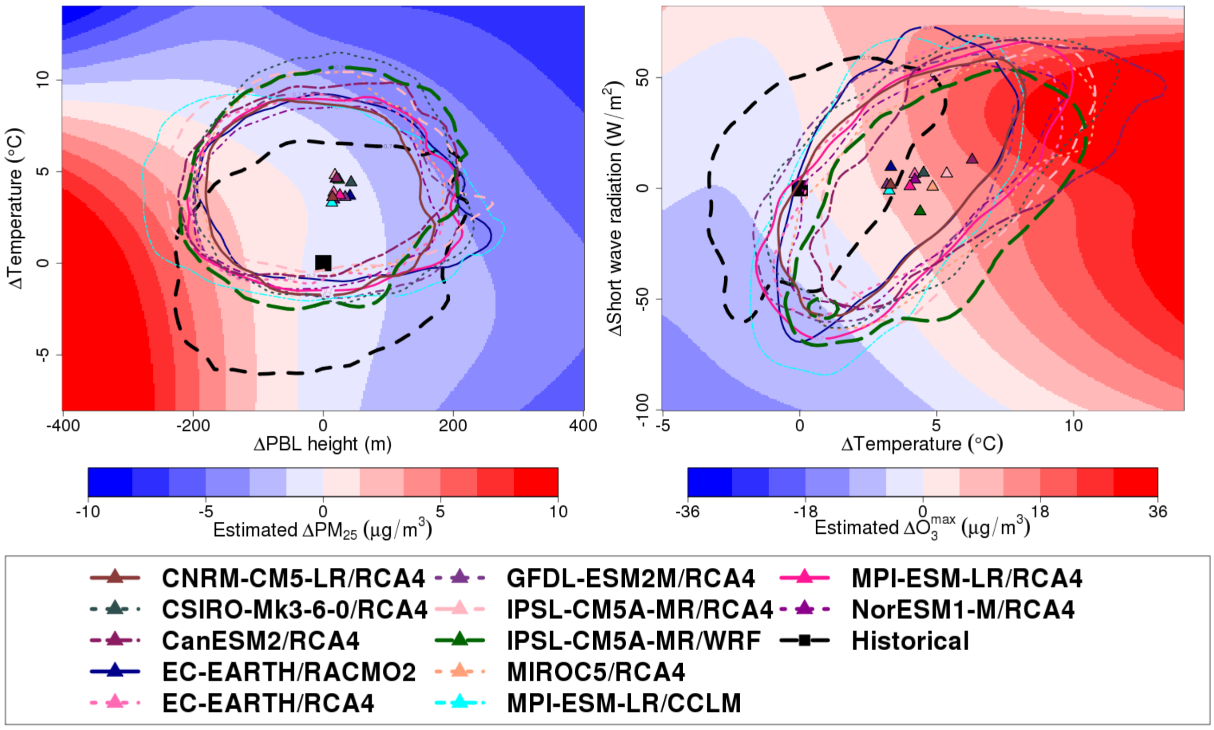

Figure 5: The left figure presents the proxy of ensemble projections for daily average de-seasonnalised PM2.5 concentrations in Eastern Europe. The right figure represents the proxy for daily maximum de-seasonnalised summer ozone for Eastern Europe. For both figures, the shaded background represents the evolution of pollutants estimated by the statistical models. The contours are representing the regional climate projections and the triangles their mean. The black dashed contour represents the historical – IPSL-CM5A-MR/WRF – and the square its mean.

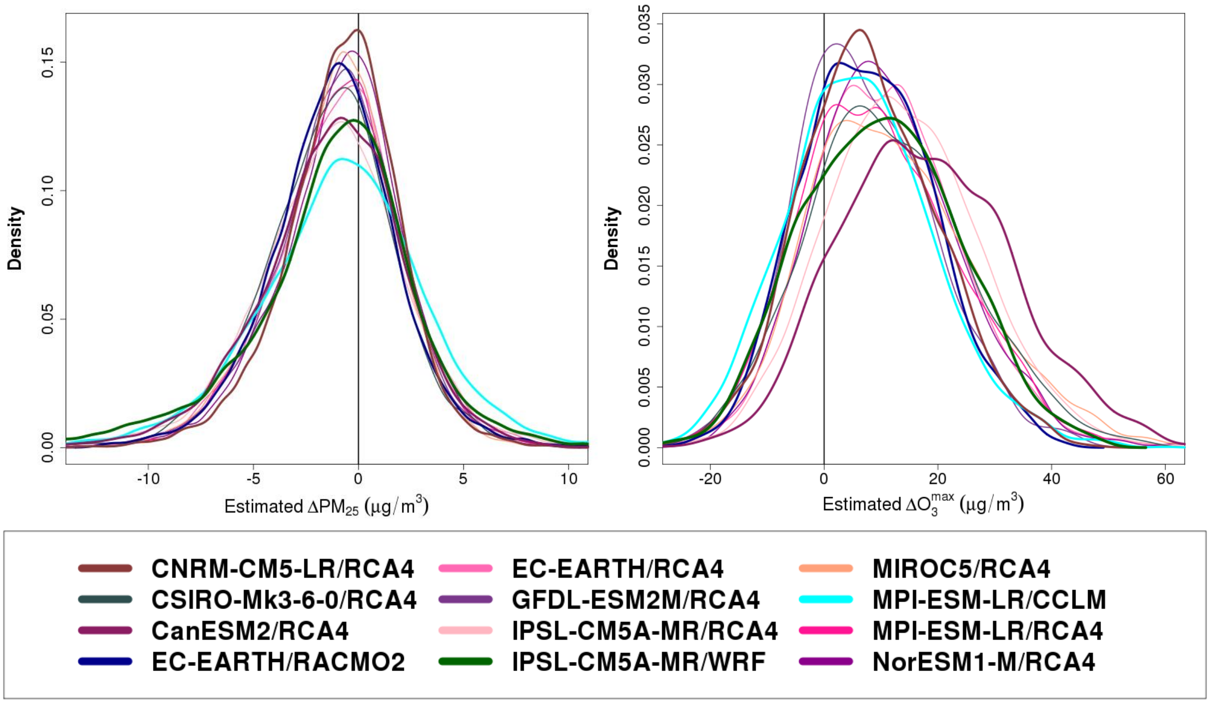

Figure 6: The left figure represents, for each regional climate model the probability density function (PDF) of the concentrations estimated with the generalized additive model at the end of the century minus the estimated concentrations of the historical period for daily average de-seasonnalised PM2.5 concentrations in Eastern Europe. The right figure presents the results for daily maximum de-seasonnalised summer ozone for Eastern Europe.

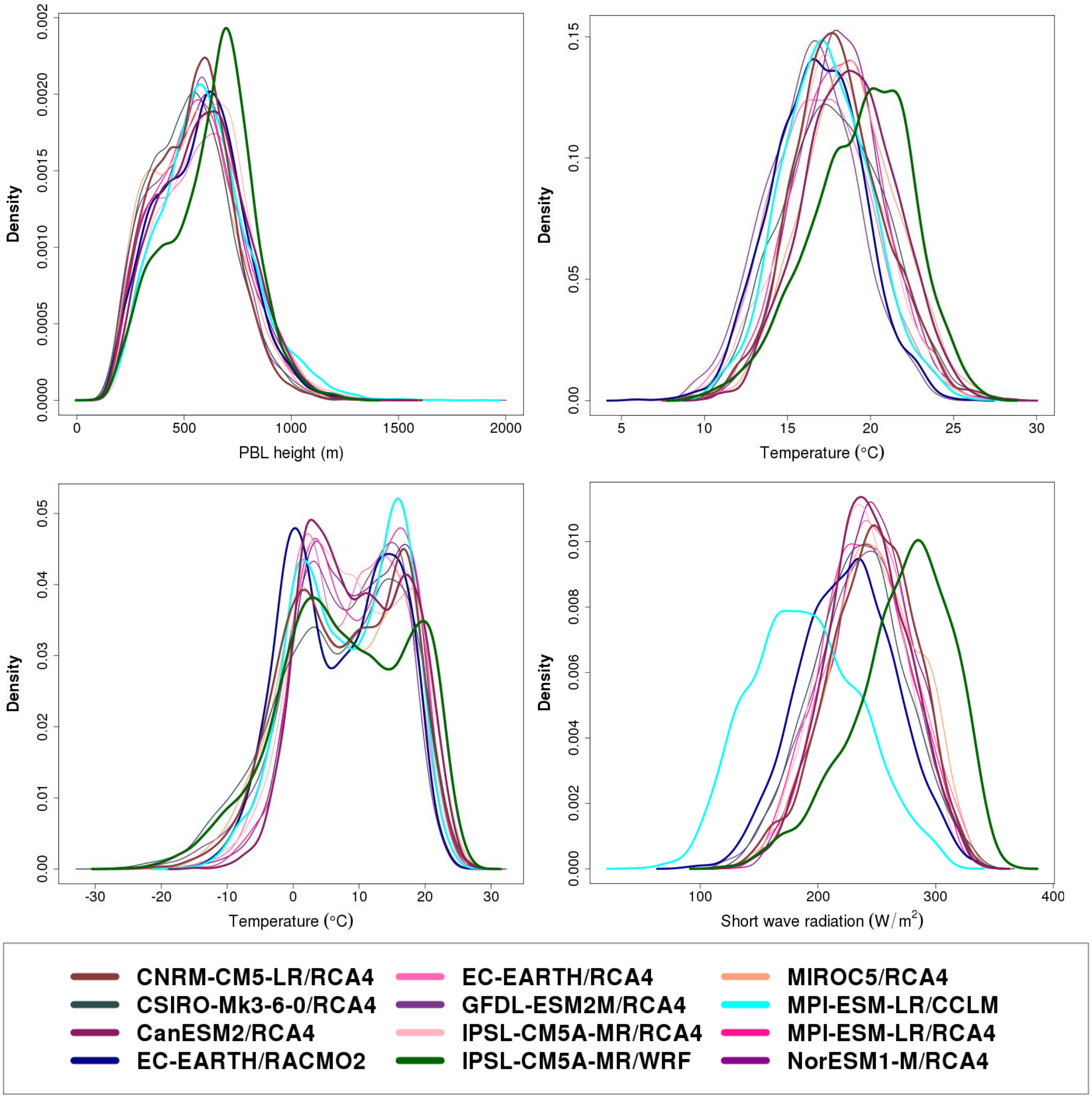

Figure 7: The first column of the panel represents the historical distribution of the meteorological variables identified by our statistical models as the two major drivers (a. PBL Height; b. near surface temperature) for PM2.5 in Eastern Europe. The second column represents the historical JJA distribution of the two main drivers for summer ozone (a. near surface temperature; b. incoming short wave radiation).